潜科学丛书

QIAN KE XUE CONG SHU

QIANKEXUE CONGSHU

数学猜想集

SHUXUE CHIXIANGJI

□ 徐本顺　解恩泽/编著

□ 湖南科学技术出版社

潜科学丛书 数学猜想集 徐本顺 解恩泽/编著 湖南科学技术出版社

潜科学丛书

数学猜想集

□ 徐本顺　解恩泽/编著

□ 湖南科学技术出版社



潜科学丛书

数学猜想集

编　　著：徐本顺　解恩泽

责任编辑：曾平安　张玉纲

出版发行：湖南科学技术出版社

社　　址：长沙市展览馆路 66 号

印　　刷：湖南省新华印刷二厂

厂　　址：邵阳市双坡岭

邮　　编：422001

(印装质量问题请直接与本厂联系)

经　　销：湖南省新华书店

出版日期：1998 年 5 月第 2 版第 2 次

开　　本：850mm×1168mm　1/32

印　　张：8.25

插　　页：4

字　　数：213000

印　　数：1301—4600

征订期号：京所科技新书目 433—228

书　　号：ISBN 7—5357—0654—1/O·73

定　　价：15.50 元

总　　序

1979年11月，在中国大地上诞生了“潜科学”这一新概念。作为一门学科，“潜科学”学一方面要研究创新性的科学技术思想胚胎从潜到显的内部孕育过程的基本规律，以寻求最大限度发挥人们科学创造潜力的途径；另一方面要研究新观点、新学说，从提出、传播、鉴别和检验进入科学殿堂的外部成长过程的基本规律，以确定使新理论顺利成长的适宜条件。作为一项事业，“潜科学”将利用刊物、年鉴、学术讨论和科学基金等多种手段，积极发掘富有开拓精神和创造才能的科技人才，热情扶持已经萌发的新思想、新学说的成长，帮助它们冲破种种障碍，为科学百花园不断增添新的奇葩，推动学术上的自由探讨和繁荣。

现代科学技术的各个部门都在加速向前发展，随着每一个领域里的惊人进步，在人们面前展现出愈来愈广阔的未知世界。传统观念和理论受到有力的冲击和挑战，层出不穷的新课题激励着

人们去探索；现代技术的突破性进展，使新技术革命的浪潮席卷全球，正在引起生产组织、产业结构和社会生活的重大变革。在这种形势下，积极推动潜科学理论的研究和潜科学事业的发展，特别是推动那些具有潜科学价值和未来意义的开发性探索，更是有特殊意义的。

为了促进这一新兴学科的成长，推动这一新生事业的发展，由“中国潜科学研究会”与《潜科学》杂志社共同组织，并系统地编写一套“潜科学丛书”。旨在通过对科学技术发展中大量个例的剖析，从不同的侧面和角度，揭示科学技术更替变革的历史足迹，概括出某些共同的带规律性的东西，以总结经验，吸取教训，为新思想、新观点、新假说、新理论的孕育和成长摇旗呐喊，鸣锣开道。

“潜科学丛书”是一套带有学术性、探索性、哲理性和趣味性的文集。我们要求每篇文章史料要翔实，科学内容要准确，观点要鲜明，力求作到文献性、科学性和思想性的统一，为进一步的深入研究提供启示。

这套丛书，自1986年以来，先后出版了《科学史上的重大争论集》、《科学蒙难集》、《科学发现个例分析》、《技术发明个列分析》、《数学猜想》、《科学前沿疑难与展望》六本。受到了国内外读者的好评，1996年获全国优秀科普读物三等奖。许多读者希望这套丛书能重新出版。为了不辜负读者的厚爱，我们将已出版的六本书作了重新修订，书名改为《科学争论集》、《科学蒙难集》、《科学发现集》、《技术发明集》、《数学猜想集》、《科学前沿集》，另外精编增补了《科学悖论集》和《科学问题集》两本，一套总共八本，奉献给读者。

当前，正是大力倡导“科教兴国”之时，这套丛书重编再版其意义更为深远，我们可以从这套丛书中，找到更多的科学技术发展的潜在规律，以促进我国科学技术的更快发展。

这套丛书的编写，是一个有益的尝试。我们希望吸引、动员

更多具有创新精神和见解的潜科学事业的支持者投入这套丛书的编写工作，不断扩大范围，丰富内容和提高质量，在推进科学技术事业的发展中，起到它的一点作用。

目　　录

序 …… (1)
一、源远流长——从勾股定理到费尔马大定理 …… (3)
(一) 从勾股三角形谈起 …… (4)
(二) 勾股数 …… (6)
(三) 问题的拓广与特例 …… (12)
1. 由"变"到"常"，由"常"到"变" …… (13)
2. 由相同到相异 …… (18)
3. 由多到少，由少到多 …… (28)
4. 由特殊到一般，由一般到特殊 …… (34)
5. 由形到数，由数到形 …… (48)
6. 从数的性质提出问题 …… (53)
7. 由类比提出问题 …… (57)
(四) 一条著名的旁注 …… (60)
二、长路漫漫——费尔马大定理的探讨 …… (63)
(一) $n=4$ 的费尔马大定理 …… (63)
(二) 关于 $n=3$ 的欧拉证明 …… (66)

1. 欧拉关于 $n=3$ 的证明 …… (67)
2. 根式环 …… (70)
3. 关于两平方数之和 …… (74)
4. 一个引理的证明 …… (81)
5. 关于两个平方数之和的注记 …… (85)
6. 欧拉证明的基本思路 …… (89)
(三) 关于 $n=3$ 的一个初等证明 …… (91)
(四) 从勒让德到库姆尔 …… (103)
1. 关于 $n=5$ 和 $n=7$，分圆整数 …… (103)
2. 代数数论基本知识 …… (105)
3. 关于正则素数 …… (110)
4. 其它一些结果 …… (111)
(五) 费尔马大定理研究的一些新成果 …… (112)
1. 考虑结论反面的必要条件 …… (112)
2. 充分条件 …… (114)
(六) 简评 …… (115)
三、触类旁通——费尔马大定理与莫德尔猜想 …… (120)
(一) 莫德尔猜想 …… (120)
(二) 解不定方程的一般性问题 …… (122)
(三) 几个重要结果 …… (121)
31. 曲线的沙发列维奇 (shafarevich) 猜想 …… (123)
2. 阿贝尔簇的沙发列维奇猜想 …… (124)
3. 有界高度原理 …… (124)
4. 同源下高的行为 …… (125)
5. 泰特猜想 …… (126)
(四) 莫德尔猜想的证明 …… (126)
(五) 从莫德尔猜想到费尔马大定理 …… (127)
(六) 模曲线和费尔马大定理 …… (129)
(七) 费尔马大定理获证之后 …… (130)
四、一步之遥？——哥德巴赫猜想 …… (131)
(一) 猜想的提出 …… (131)
(二) 悲观的预言与惊人的成果 …… (133)
(三) 圆法 …… (135)

（四）筛法 …………………………………………………………… (145)
五、补天何须五色石——地图着色与四色猜想 ……………… (152)
（一）四色猜想的提出 ………………………………………… (152)
1. 什么叫四色猜想 …………………………………………… (152)
2. 先生问学生和学生问先生 ………………………………… (153)
（二）早期的证明和五色定理 ………………………………… (154)
1. 凯利的呼吁 ………………………………………………… (154)
2. 另辟蹊径 …………………………………………………… (155)
3. 约当曲线和欧拉定理 ……………………………………… (156)
4. 五色定理 …………………………………………………… (158)
5. 肯普的证明 ………………………………………………… (160)
（三）四色猜想的证明 ………………………………………… (161)
1. 不可避免组和可约构形 …………………………………… (161)
2. 公开宣称的一种信念 ……………………………………… (162)
3. 等价的形式 ………………………………………………… (163)
4. 可约性障碍和放电 ………………………………………… (164)
5. 新的困难 …………………………………………………… (164)
6. 人机合作证明了四色猜想 ………………………………… (165)
7. 解决地图四色问题的重大意义 …………………………… (166)
（四）平面图 …………………………………………………… (166)
（五）线（边）着色 …………………………………………… (168)
（六）顶点着色 ………………………………………………… (173)
（七）全色猜想 ………………………………………………… (178)
六、法无定法——提出数学猜想的若干方法 ………………… (180)
（一）不完全归纳法 …………………………………………… (180)
（二）类比法 …………………………………………………… (184)
（三）变换条件法 ……………………………………………… (187)
（四）物理模拟法 ……………………………………………… (187)
（五）联系观察法 ……………………………………………… (188)
（六）逐级猜想法 ……………………………………………… (192)
七、闪光的并非都是金子——判定数学猜想真伪性的几个途径 ……………………………………………………………… (194)
（一）举例否定 ………………………………………………… (194)

（二）逐次趋近 …… (197)
（三）命题转化 …… (200)
（四）反证法 …… (201)
八、千淘万漉始到金——数学猜想的艰难性 …… (206)
（一）有一个逐步完善的过程 …… (206)
（二）时间长与途径曲折 …… (208)
1. 时间长 …… (208)
2. 途径曲折 …… (209)
（三）有时得不到多数人的承认 …… (214)
九、数学猜想的类型、特征与意义 …… (218)
（一）数学猜想的类型 …… (218)
1. 存在型猜想 …… (218)
2. 规律型猜想 …… (219)
3. 方法型猜想 …… (220)
（二）数学猜想的特征 …… (221)
1. 真伪的待定性 …… (221)
2. 思想的创新性 …… (222)
3. 目标的具体性 …… (223)
（三）研讨数学猜想的重要意义 …… (224)
1. 丰富数学理论 …… (225)
2. 促进数学方法论的研究 …… (227)
3. 推动潜科学学的探讨 …… (230)

序

加强潜科学的理论研究，不仅要结合历史资料对潜科学的基本形态作概括性的分析，而且更要对其基本形态中的具体表现作深入的探讨，从中发现规律性的东西，使潜科学学的学科体系建立在更坚实的基础之上。这本书，正是本着这样的想法写成的。

所谓数学猜想，是指根据某些已知的事实材料和数学知识，对未知的量及其关系所作出的一种预测性的推断。它既有一定的科学性，又有某种假定性，是科学性与假定性的辩证统一。一般说来，这种推断是数学中难度较大的问题。数学猜想不仅是数学研究的一个科学方法和数学发展的一种重要形式，而且还是科学猜测这种潜科学基本形态在数学中的具体表现。因此，对它进行考察和分析，必将促进潜科学理论研究的深入开展。

本书的重点是分析和阐述数学猜想的思想与方法。全书共有九个部分，从内容上可分为两大类：第一类是对数学猜想的典型事例进行历史考察与思想剖析（一、二、三、四、五）；第二类

是就数学猜想的产生、演变、判定、类型、特征和意义等，作综合性的分析和论述（六、七、八、九）。为了更系统地了解数学发展史上著名数学猜想的研究状况，我们特编制了《数学猜想一览表》，作为〔附录〕载于书后，仅供参考。在写作中，我们力图资料准确、史实具体、观点鲜明、重点突出，在揭示“潜”的实质与规律性上下功夫。

数学猜想大都是专业性很强的数学难题，写作中遇到的最大困难就是通俗化的问题。我们虽作了很大的努力，但也难满足广大读者的需要。

在写作过程中，我们参考和引述了国内外许多重要文献。特别是华罗庚、王元、陈景润、柯召、孙琦、潘承洞、潘承彪、闵嗣鹤、刘彦佩、冯克勤等我国著名数学家和学者的有关论著，对我们有很大的启发和帮助。

李迪教授对本书的写作给予了许多支持，姚俊梅、徐炎章参加了本书的部分编写工作，这里，一并致谢。

写这样的书，对我们来说还是一个尝试。限于我们的水平及问题本身的难度，不妥之处在所难免，敬请读者批评指正。

徐本顺　解恩泽

一、源远流长

——从勾股定理到费尔马大定理

17世纪，法国数学家彼埃尔·德·费尔马(Pierre de Fermat，1601—1665)曾宣称他证明了：当整数 n 大于2时，不存在一个整数 n 次幂是另外两个整数 n 次幂之和。三个世纪以来，虽然许许多多数学家试图证明它，但始终没有成功。费尔马当时是一位法官，并非职业数学家，居然以他的“定理”难住了三个世纪以来的优秀数学家，这种思想真是太微妙了！这个问题看来似乎很简单，只要有初中水平的数学知识，就不难理解所提出的问题，但为了证明它，却吸引了千千万万个数学爱好者。他们一代又一代，前赴后继，有的甚至为之献出毕生精力，但均未达到最终目的。这个困惑人们三百多年的“大定理”，直至1993年才被英国数学家安德鲁·维尔斯攻克。我们简略地回顾一下人们研讨这个问题的历史过程及其所获得的成果，无疑对我们进一步解决这个问题，以及增强提出问题和解决问题的能力是会有一定的后发作用的。

（一）从勾股三角形谈起

早在3500年以前，巴比伦人就知道三边为下列各数的一些三角形为直角三角形：

120,	119,	169;
3456,	3367,	4825;
4800,	4601,	6649;
13500,	12709,	18541;
72,	65,	97;
360,	319,	481;
2700,	2291,	3541;
960,	799,	1249;
600,	481,	769;
6480,	4961,	8161;
60,	45,	75;
2400,	1679,	2929;
240,	161,	289;
2700,	1771,	3229;
90,	56,	1060。

然而，当时为什么列出这些三角形，现在还是个谜。

在3000年以前，中国已经知道用边长为3，4，5的直角三角形来进行测量。中国最古老的一本天文数学书《周髀算经》一开头就有周公问商高："天不可阶而升，地不可得尺寸而度。"商高讲："勾广三，股修四，经隅五。"① 其意是天的高度和地面上的一些测量的数字是怎样得到的呢？这可用边长为3，4，5的直角三角形来测算。在《周髀算经》卷上之二，荣方、陈子问答

① 钱宝琮校点，算经十书（上册），中华书局，1963年，第13页。

中，陈子讲道："勾股各自乘，并而开方除之"，[①] 用现代的数学符号表示，就是

$$a^2+b^2=c^2,$$

其中，a，b 是三角形的直角边，c 为其斜边。可见，陈子已经知道一般的勾股定理。约公元 222 年，赵君卿为《周髀算经》注释时，对勾股定理作了理论证明。

勾股定理，西方称为毕达哥拉斯定理。毕达哥拉斯(Pythagoras，古希腊人，约前 572—前 500 年）于公元前 572 年左右出生在爱琴海的萨摩斯岛。据说，他从师于泰勒斯，可能还游历过埃及、巴比伦。他在意大利南部的希腊油港克罗托内创办了著名的毕达哥拉斯学校。这个学校不仅研究哲学、数学，而且发展成了一个有秘密仪式和盟约的帮会式的团体。由于该团体除了搞学术活动外，还参与政治活动，同当时的贵族党结成联盟，因此南意大利的民主力量摧毁了学校的建筑并迫使该团体解散。据传，毕达哥拉斯于公元前 500 年左右被害于梅塔庞通。他的弟子则分散到其他的学术中心去了。该团体虽然形式上不存在了，但这个学派实际上继续存在至少两个世纪之久。该学派的学术思想和成就影响十分深远。由于该团体所有发现都归功于毕达哥拉斯，因此现在很难确切知道那项发现应归功于学派中哪一个人了，毕达哥拉斯的学术成就是集体智慧的产物。

"万物皆数"[②] 是毕达哥拉斯学派对宇宙和事物的本质认识。他们认为整数是人和物质的各种各样性质的起因。在这种哲学思想的指导下，就导致他们对于数的性质的阐述和探讨。因此，把几何问题与算术紧密结合起来研究，就成为毕达哥拉斯学派的很自然的事情了。3，4，5 为边的直角三角形三边皆为整数，这种三条边为整数的直角三角形称为勾股三角形，或者叫做毕达哥拉

---

① 钱宝琮校点，算经十书（上册），中华书局，1963 年，第 27 页。

② M·克莱因著，张理京、张锦炎译，《古今数学思想》（第一册），上海科学技术出版社，1979 年，第 168 页。

斯三角形。把所有的勾股三角形都找出来就等价于求下列不定方程

$$x^2+y^2=z^2$$

的所有正整数解，于是，一个几何问题就转化为一个数论中的不定方程求解问题。

（二）勾股数

如果正整数 x，y，z 能满足下列不定方程

$$x^2+y^2=z^2, \tag{1}$$

则它们叫做勾股数。

求出（1）的所有正整数解，也就是求出（1）的所有勾股数。

为了解决这个初等数论中的著名问题，我们再观察几个简单的勾股三角形：

3,	4,	5;
5,	12,	13;
7,	24,	25;
9,	40,	41;
11,	60,	61;
……		

观察这些数，可发现如下规律：

第一个数是奇数，第二个数是第一个数的平方减 1 再除以 2，第三个数是第二个数加 1，也是第一个数的平方加 1 再除以 2，即设 m 为奇数，则一般有

$$m,\ \frac{m^2-1}{2},\ \frac{m^2+1}{2}。$$

于是有

$$m^2+\left(\frac{m^2-1}{2}\right)^2=\left(\frac{m^2+1}{2}\right)^2, \tag{2}$$

其中 m 为奇数。

一般认为毕达哥拉斯已发现了公式（2）[1] 不过公式（2）只给出一部分勾股数。

（2）式两边同乘上 4，再变形，得

$$(2m)^2+(m^2-1)^2=(m^2+1)^2。\tag{3}$$

显然（3）式不论 m 是奇数还是偶数，等式都成立。古希腊哲学家、数学家柏拉图（Plato，约公元前 430—约公元前 349），于公元前 380 年给出了公式（3），其中 m 为奇数或偶数。

显然公式（2）与（3）均不能给出全部的勾股三角形。实际上，如果以 a，b，c 为边的直角三边形为勾股三角形，那么，对于任意自然数 n，以 na，nb，nc 为边的三角形也必定是勾股三角形。

在公式（3）中，m 为任意自然数，1 是一个特殊的自然数，由特殊到一般，若 1 也变成任意自然数，比如变成 n^2，是否能给出所有的勾股三角形呢？把 1 变成 n^2，为了使（3）式保持恒等，（3）中的第一项$(2m)^2$ 应变成$(2mn)^2$，即有

$$(2mn)^2+(m^2-n^2)^2=(m^2+n^2)^2。$$

显然，当 $m=n$ 时，就不能得到三角形，更谈不上勾股三角形了，因此，需要对 m，n 作些限制。

大约公元 3 世纪，亚历山大里亚的丢番图（Diophantus，约 246—330）给出一个方法，可得到全部的勾股三角形。具体可叙述如下。

如果（1）式有正整数解，且满足$(x,y)=d>1$，则由 $d^2\mid x^2$[2]，$d^2\mid y^2$，得

$$d^2\mid(x^2+y^2)。$$

而 $x^2+y^2=z^2$，所以 $d^2\mid z^2$。故有

---

① M·克莱因著，张理京，张锦炎译：《古今数学思想》（第一册），上海科学技术出版社，1979 年，第 36 页。

② 设 a，b 为二整数且 $b\neq0$，若 b 整除 a，记为 $b\mid a$；b 不整除 a，记为 $b\nmid a$。

$d \mid z$。

这时，可将（1）式两边同时约去 d。由于

$(x/d, y/d)=1$，所以在(1)中，不妨假定$(x,y)=1$。这就是说，我们先求出 $x^2+y^2=z^2$ 满足$(x,y)=1$的所有解，然后乘上一个适当因子，就能求出 $x^2+y^2=z^2$ 的所有解。

如果（1）式有正整数解，且满足$(x,y)=1$，则由$(x,y)=1$得知，x，y 不可能都是偶数，也不可能全为奇数。假定它们全是奇数，则

$$x^2 \equiv 1(\bmod 4)\text{①}, y^2 \equiv 1(\bmod 4)，就有$$

$$z^2=x^2+y^2 \equiv 2 \ (\bmod 4)。$$

这是不可能的。因此，如果（1）式有正整数解，且满足$(x,y)=1$，则 x，y 中有一个是偶数，另一个为奇数，不妨假定 x 为偶数。于是我们可给出如下定理。

定理 1.2.1 不定方程（1）的适合条件

$$x>0,\ y>0,\ z>0,\ (x,y)=1, 2 \mid x \tag{4}$$

的一切正整数解的充分必要条件是

$$x=2mn,\ y=m^2-n^2,\ z=m^2+n^2, \tag{5}$$

其中 m，n 都是正整数，且有 $m>n$，$(m,n)=1, m \not\equiv n(\bmod 2)$。

证法 1 充分性。因为 m，n 都是正整数，且 $m>n$，所以，由（5）式，得

$x=2mn>0$，$y=m^2-n^2>0$，$z=m^2+n^2>0$，即 x，y，z 满足（4）中的条件 $x>0$，$y>0$，$z>0$，$2 \mid x$。由（5）式，我们有

$$x^2+y^2=(2mn)^2+(m^2-n^2)^2=m^4+2m^2n^2+n^4$$
$$=(m^2+n^2)^2=z^2。$$

---

① 若 $m \mid a-b$，记为 $a \equiv b(\bmod m)$；
若 $m \nmid a-b$，记为 $a \not\equiv b(\bmod m)$。

所以（5）式满足（1）式的正整数解。

再证明也满足（4）中的条件$(x,y)=1$。

为此设$(x,y)=d$，若能证明$d=1$，则充分性得证。事实上，我们有$d|x$，$d|y$，$d^2|x^2$，$d^2|y^2$，故有

$$d^2\mid(x^2+y^2)。$$

而 $x^2+y^2=z^2$，即$d^2|z^2$，故$d|z$。设$y=dl$，$z=dk$，则由（5）式，我们有

$$2m^2=m^2-n^2+m^2+n^2=y+z=d(l+k),$$

$$2n^2=m^2+n^2-(m^2-n^2)=z-y=d(l-k)。$$

所以有$d|2m^2$，$d|2n^2$，故有$d\mid 2(m^2,n^2)$。由于在（5）中有$(m,n)=1$，故有$(m^2,n^2)=1$，从而知$d|2$。由于$m-n=m+n-2n$和（5）中有$m\not\equiv n(\bmod 2)$，知

$$2\nmid(m+n),2\nmid(m-n),$$

即 $2\nmid(m^2-n^2)$。又由(5)知$2\nmid y$。综合$2\nmid y$，$d\mid 2$和$(x,y)=d$，得知$d=1$，即$(x,y)=1$。

由上可知，（5）式满足（4）中所有条件，从而充分性得证。

必要性。若x，y，z是适合（4）中的所有条件的（1）式的任一组解时，则x，y，z一定具有（5）的形式。

我们先证明$\frac{z+y}{2}$和$\frac{z-y}{2}$都是整数，且互素。因x是偶数，y为奇数，所以x^2+y^2必为奇数，而$x^2+y^2=z^2$，故知z^2为奇数。于是z也是奇数。由于z，y都是奇数，所以$z+y$和$z-y$都是偶数，因而$\frac{z+y}{2}$和$\frac{z-y}{2}$都是整数。为了证明它们互素，我们假定

$\left(\frac{z+y}{2},\frac{z-y}{2}\right)=d$，再证$d=1$。将$\left(\frac{z+y}{2},\frac{z-y}{2}\right)=d$改写成

$$\frac{z+y}{2}=l_1d, \tag{6}$$

$$\frac{z-y}{2}=l_2d, \tag{7}$$

其中$(l_1, l_2) = 1$。由（6）－（7），得

$$y = (l_1 - l_2)d。$$

所以知$d \mid y$。又由（1）知

$$x^2 = z^2 - y^2 = (z + y)(z - y) = 4l_1l_2d^2,$$

故得　$d^2 \mid x^2$，即$d \mid x$。于是有$d \mid (x, y)$。而$(x, y) = 1$，故$d = 1$。从而得证$\frac{z+y}{2}$和$\frac{z-y}{2}$互素。

将（1）改写成

$$x^2 = (z + y)(z - y)。 \tag{8}$$

（8）式两边同除以4，得

$$\left(\frac{x}{2}\right)^2 = \left(\frac{z+y}{2}\right)\left(\frac{z-y}{2}\right)。$$

由于$2 \mid x$，所以上式左端是一个平方数。由于$\left(\frac{z+y}{2}, \frac{z-y}{2}\right) = 1$，

所以知$\frac{z+y}{2}$，$\frac{z-y}{2}$一定都是平方数。令

$$\frac{z+y}{2} = m^2,$$

$$\frac{z-y}{2} = n^2。 \tag{9}$$

由于x，y，z都是正整数，故知$m > n > 0$。由$\left(\frac{z+y}{2}, \frac{z-y}{2}\right) = 1$，知$(m^2, n^2) = 1$，因而$(m, n) = 1$。（9）式中的二式相加相减分别得到

$$z = m^2 + n^2,$$

$$y = m^2 - n^2。$$

再由（8），（9）得到$x^2 = 4m^2n^2$，从而

$$x = 2mn。$$

又由于z是奇数，而$z = m^2 + n^2$，所以m，n中必定有一个为偶数，另一个为奇数。于是$m + n$是奇数，所以

$m \not\equiv n \pmod 2$。（证完）

上述的定理是直接用初等数论的方法给予证明的。我们也可以将（1）式变形，利用变量替换，转化成三角函数，再用三角函数的有关性质给出另外的证明方法。为此我们先证一个引理。

引理 $\sin\alpha$ 和 $\cos\alpha$ 都为有理数的充分必要条件是 $\operatorname{tg}\frac{\alpha}{2}$ 为有理数或者无意义。

证明 先证充分性。由

$$\sin\alpha = \frac{2\operatorname{tg}\frac{\alpha}{2}}{1+\operatorname{tg}^2\frac{\alpha}{2}},$$

$$\cos\alpha = \frac{1-\operatorname{tg}^2\frac{\alpha}{2}}{1+\operatorname{tg}^2\frac{\alpha}{2}}$$

易知，当 $\operatorname{tg}\frac{\alpha}{2}$ 为有理数时，$\sin\alpha$，$\cos\alpha$ 都为有理数。若 $\operatorname{tg}\frac{\alpha}{2}$ 无意义，则

$$\frac{\alpha}{2} = k\pi + \frac{\pi}{2}，\ k \text{ 为整数},$$

即 $\alpha = (2k+1)\pi$。

所以 $\sin\alpha = 0$，$\cos\alpha = -1$。

再证必要性。由半角公式

$$\operatorname{tg}\frac{\alpha}{2} = \frac{\sin\alpha}{1+\cos\alpha}$$

易知，当 $\sin\alpha$，$\cos\alpha$ 为有理数时，$\operatorname{tg}\frac{\alpha}{2}$ 为有理数。若 $\alpha = (2k+1)\pi$；$\sin(2k+1)\pi = 0, \cos(2k+1)\pi = -1$，则 $\operatorname{tg}\frac{\alpha}{2}$ 无意义。

综上所述，引理得证。

证法 2 将（1）式变形，得

$$\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 = 1。\tag{10}$$

令 $\frac{x}{z}=\sin\alpha$， $\frac{y}{z}=\cos\alpha$。因为 $x>0$，$y>0$，$z>0$，所以，$\sin\alpha$，$\cos\alpha$ 为正有理数。不失一般性，假定 $0<\alpha<\frac{\pi}{2}$，则 $0<\frac{\alpha}{2}<\frac{\pi}{4}$。

由万能置换公式

$$\sin\alpha=\frac{2\mathrm{tg}\frac{\alpha}{2}}{1+\mathrm{tg}^2\frac{\alpha}{2}},$$

$$\cos\alpha=\frac{1-\mathrm{tg}^2\frac{\alpha}{2}}{1+\mathrm{tg}^2\frac{\alpha}{2}}$$

和引理，可令 $\mathrm{tg}\frac{\alpha}{2}=\frac{n}{m},(m,n)=1$。

由上述条件有 $m>n>0$，则

$$\frac{x}{z}=\frac{2\cdot\frac{n}{m}}{1+\left(\frac{n}{m}\right)^2}=\frac{2mn}{m^2+n^2}, \tag{11}$$

$$\frac{y}{z}=\frac{1-\left(\frac{n}{m}\right)^2}{1+\left(\frac{n}{m}\right)^2}=\frac{m^2-n^2}{m^2+n^2}。 \tag{12}$$

由（11）÷（12），得

$$\frac{x}{y}=\frac{2mn}{m^2-n^2}。$$

由$(x,y)=1$，可知 m 与 n 中必有一奇一偶，且 $x=2mn$，$y=m^2-n^2$，$z=m^2+n^2$。（证完）

（三）问题的拓广与特例

旧的问题解决了，新的问题又会产生。在数学体系内部，通

过一般化、特殊化和各种各样的逻辑组合等方式，可以不断地向自身提出新的问题。从不同的角度去看§2 中的（1）式，还可以提出各种各样的问题。这些问题的提出，需要我们去分析解决。那么，由(1)式可提出哪些问题呢？本节将做较详细的叙述。

1 由“变”到“常”，由“常”到“变”

在（二）中的（1）式，底数为变量，指数为常量 2。若底数分别取特殊的常量，让指数成为变量，比如 x，y，z 分别取 3，4，5，就可提出：

求不定方程

$$3^x+4^y=5^z$$

的所有正整数解的问题。

通过试验，除了 $x=y=z=2$ 外，似乎还找不出其它正整数解适合此方程，那么是否该方程只有这一组正整数解呢？答案是肯定的，这可从理论上进行证明。于是有

定理 1.3.1 不定方程

$$3^x+4^y=5^z \tag{1}$$

有且仅有正整数解 $x=y=z=2$。

证明 由（1），得

$$(4-1)^x+4^y=(4+1)^z, \tag{2}$$

或

$$3^x+(3+1)^y=(6-1)^z。 \tag{3}$$

由（2）及（3）和牛顿二项式定理，得知

$$(-1)^x\equiv 1(\bmod 4),$$

$$1\equiv(-1)^z(\bmod 3)。$$

因此 x 和 z 均为偶数，设 $x=2a$，$z=2b$，（1）式就化为

$$(3^a)^2+(2^y)^2=(5^b)^2。 \tag{4}$$

由勾股定理知

$$2^y=2AB,\ 3^a=A^2-B^2,\ 5^b=A^2+B^2,$$

且 $A>B>0$，$(A,B)=1$，A，B 中一奇一偶。显然 $B=1$，从而 $A=2^{y-1}$，于是

$$3^a=(2^{y-1})^2-1=(2^{y-1}+1)(2^{y-1}-1)。$$

因为 3 是奇数，且知 $2^{y-1}+1>2^{y-1}-1$，$(2^{y-1}+1,2^{y-1}-1)=1$，故

$$2^{y-1}-1=1，\ 2^{y-1}+1=3^a。$$

于是有

$$2^{y-1}=2。$$

解上述方程，得

$$y=2。$$

又由 $2^{2-1}+1=3=3^a$，得

$$a=1。$$

再由 $5^b=(2^{y-1})^2+1=2^2+1=5$，得

$$b=1。$$

综上所述，命题得证。 （证完）

在方程

$$3^x+4^y=5^z$$

中，3，4，5 是固定的勾股数。如果让固定的勾股数也变动起来，即已知 a，b，c，且 $a^2+b^2=c^2$，那么方程

$$a^x+b^y=c^z$$

会有怎样的结论呢？换言之，对于正整数 a，b，c，x，y，z，如果有

$$a^2+b^2=c^2$$

和

$$a^x+b^y=c^z，$$

求关于 x，y，z 的指数不定方程的正整数解。

为了解决这一问题，可取一些特殊勾股数进行考察。我们已经对勾股数 3，4，5 得到了结论是 $x=y=z=2$，我们再取一些特殊的勾股数，比如取

5，　　　　12，　　　　13；

7,	24,	25;
9,	40,	41;
11,	60,	61。

经研究，也有同样的结论：

$$x = y = z = 2。$$

一般地，可提出如下猜想：

对于正整数 a，b，c，x，y，z，如果有

$$a^2 + b^2 = c^2$$

和 $$a^x + b^y = c^z,$$

那么 $$x = y = z = 2。$$

这一猜想至今没有解决。但对于一些特殊 a，b，c 得到如下一些结果。

当 $a = 2n + 1,\ b = 2n(n+1),\ c = 2n(n+1) + 1,$
$$n > 0 \tag{5}$$
时，有

定理 1.3.2 对于（5）中的数，当

$$n \equiv 1, 4, 5, 9, 10 (\mathrm{mod} 12)$$

时，猜想均成立。

定理 1.3.3 对于（5）中的数，如果

i) $n \equiv 1(\mathrm{mod} 2)$，且有素数 p 存在，使得

$$2n + 1 = p^s,\quad s > 0,$$

或

ii) $n \not\equiv 3(\mathrm{mod} 4)$，且有素数 $p \equiv 3(\mathrm{mod} 4)$ 存在，使得

$$2n + 1 \equiv 0(\mathrm{mod}\, p),$$

则猜想成立。

定理 1.3.4 对于（5）中的数，$n < 6144$ 时，猜想成立。

定理 1.3.5 对于勾股数

$$a = 4n^2 - 1,\quad b = 4n,\quad c = 4n^2 + 1, \tag{6}$$

上述猜想成立。

证明 当 $n=1$ 时，就是定理1.3.1。故可设 $n>1$。

先证 x，z 为偶数，$y\geqslant 2$，再证 $x=y=z=2$。将（6）代入 $a^2+b^2=c^2$ 中，得

$$(4n^2-1)^2+(4n)^2=(4n^2+1)^2, \tag{7}$$

从而由牛顿二项式定理，可推知

$$(-1)^x\equiv 1(\bmod 4),$$

于是可知 x 必为偶数。

又由（7）可知

$$(-1)^x+(4n)^y\equiv 1(\bmod 4n^2),$$

而 x 为偶数，于是有

$$(4n)^y\equiv 0(\bmod 4n^2)。$$

因为 $n>1$，所以推知 $y\geqslant 2$。

由（6）可得

$$a^2=16n^4-8n^2+1,$$
$$c^2=16n^4+8n^2+1,$$

故推知

$$a^2\equiv 1(\bmod 8n^2),$$
$$c^2\equiv 1(\bmod 8n^2)。$$

由此可知，z 必为偶数。如若不然，即设 z 为奇数，因 x 为偶数，$y\geqslant 2$，由（7）得

$$1\equiv 4n^2+1(\bmod 8n^2),$$

这是不可能的，所以 z 必为偶数。

既然 x，z 为偶数，我们可设 $x=2x_1$，$z=2z_1$。代入方程（7）中，可得

$$[(4n^2+1)^{z_1}+(4n^2-1)^{x_1}][(4n^2+1)^{z_1}-(4n^2-1)^{x_1}]$$
$$<(4n)^y。 \tag{8}$$

下面我们证明 x_1，z_1 为奇数。如若不然，设 x_1 为偶数，那么

$$(4n^2+1)^{z_1}+(4n^2-1)^{x_1}\equiv 2(\bmod 4n),$$

$$(4n^2+1)^{z_1}-(4n^2-1)^{x_1}\equiv 0(\bmod 4n)。$$

由（8），得出

$$(4n^2+1)^{z_1}+(4n^2-1)^{x_1}=2,$$

$$(4n^2+1)^{z_1}-(4n^2-1)^{x_1}=2\cdot 4^{y-1}n^y,$$

这是不可能的，所以 x_1 是奇数，这时

$$(4n^2+1)^{z_1}+(4n^2-1)^{x_1}\equiv 0(\bmod 4n),$$

$$(4n^2+1)^{z_1}-(4n^2-1)^{x_1}\equiv 2(\bmod 4n)。$$

再由（8），得出

$$(4n^2+1)^{z_1}+(4n^2-1)^{x_1}=2\cdot 4^{y-1}n^y, y>1, \tag{9}$$

$$(4n^2+1)^{z_1}-(4n^2-1)^{x_1}=2。\tag{10}$$

如果 z_1 是偶数，因 x_1 是奇数，由（10）得

$$1-(4n^2-1)\equiv 2(\bmod 8n^2),$$

这是不可能的，故 z_1 必为奇数。

（9）+（10），得

$$(4n^2+1)^{z_1}=1+4^{y-1}n^y, y>1。\tag{11}$$

下面证 $y=z$。如果 $y\geqslant 3$，因 z_1 为奇数，由（11），得

$$4n^2+1\equiv 1(\bmod 8n^2), \tag{12}$$

这是不可能的，所以必有 $y=2$。

将 $y=2$ 代入（11），得

$$(4n^2+1)^{z_1}=1+4n^2,$$

所以 $z_1=1$，即 $z=2$。

再由（10），得

$$4n^2+1-(4n^2-1)^{x_1}=2,$$

故 $x_1=1$，即 $x=2$。（证完）

在一些含有常量与变量的数学问题中，由于常量与变量所处的位置不一样，就形成了不同类型的问题。因此，当我们考虑其中一种类型的问题之后，从数学形式把常量与变量交换一下位置，就可提出另一类新型的数学问题。上面提出的指数不定方程的求解问题是由三元二次不定方程通过交换常量与变量的位置而

提出来的。有时不交换位置，将常量变量化，或将变量常量化也可提出新的问题。上述的猜想是将指数方程中的常量3，4，5变量化，且使这些变量的关系保持一个常量，即 $a^2+b^2-c^2=0$ 而提出的一类新问题。从常量和变量的相互关系中提出问题是从数学内部提出问题的常用方法之一。

2. 由相同到相异

在方程 $x^2+y^2=z^2$ 中，有元数、次数和系数之分。这是一个三元二次不定方程，每个元前面的系数都是1，若其中有的元前面的系数不是1，又会怎样呢？系数不一样，最简单的情况就是讨论不定方程

$$2x^2+y^2=z^2$$

的正整数解的问题。

怎样求该方程的正整数解呢？我们首先给这个方程正整数解的一个猜测，然后给出证明。此方程解的形式我们可仿照系数是一样情况进行如下猜测：

若 $x=2mn$，那么 y，z 的形式又该怎样呢？由 $2x^2+y^2=2(2mn)^2+y^2=8m^2n^2+y^2$ 知，应有

$$y=2m^2-n^2,\quad z=2m^2+n^2$$

或者是

$$y=m^2-2n^2,\quad z=m^2+2n^2。$$

经理论证明，它的解的形式正是这样，于是有

定理 1.3.6　不定方程

$$2x^2+y^2=z^2 \tag{1}$$

满足

$$(x,y)=1,x>0,y>0,z>0,2\mid x \tag{2}$$

的全部正整数解可表为

$$x=2mn,y=|2m^2-n^2|,z=2m^2+n^2, \tag{3}$$

其中 m，n 为整数，$m>0$，$n>0$，$2\nmid n$，且满足 $(m,n)=1$。 (4)

证明 设 m，n 满足(4)，将(3)代入(1)，得

$$2x^2+y^2=2(2mn)^2+(2m^2-n^2)^2$$
$$=(2m^2+n^2)^2=z^2。$$

显然 $x>0$，$y>0$，$z>0$。下面证 $(x,y)=1$，为此设 $(x,y)=d$，再证 $d=1$。由 $(x,y)=d$，我们有 $d\mid 2m^2-n^2, d\mid 2m^2+n^2$，于是可推出 $d\mid 2(2m^2,n^2)$，而 $(m,n)=1$，$2\nmid n$，故 $d=1$。这就证明了由(3)，(4)给出的 x，y 为(1)满足(2)的解。

反之，若(1)满足(2)的任一组解 x，y，z，一定可由(3)与(4)表出。事实上，由 $2\mid x$，故 $2\nmid y$，$2\nmid z$，由于 $(y+z,z-y)\mid 2(y,z)$，而 $(y,z)=1$，故 $(y+z,z-y)=2$，从而 $\left(y+z,\dfrac{z-y}{2}\right)=1$ 或 $\left(\dfrac{y+z}{2},z-y\right)=1$。由(1)，得

$$x^2=\frac{z+y}{2}(z-y) \text{ 或 } x^2=(z+y)\left(\frac{z-y}{2}\right),$$

故 $\quad z+y=4m^2,\ \dfrac{z-y}{2}=n^2,\ 2\nmid n,\ x=2mn,$

$\quad m>0, n>0, (m,n)=1,$

或 $\quad \dfrac{z+y}{2}=m^2,\ z-y=4n^2,\ 2\nmid n,\ x=2mn,$

$\quad m>0,\ n>0,\ (m,n)=1,$

即得 $\quad y=2m^2-n^2,\ x=2mn,\ z=2m^2+n^2,$

$\quad (m,n)=1,\ m>0,\ n>0,\ 2m^2>n^2。$

或 $\quad y=n^2-2m^2,\ x=2mn,\ z=2m^2+n^2,$

$\quad (m,n)=1,\ m>0,\ n>0,\ n^2>2m^2。$(证完)

讨论了(1)之后，我们再讨论不定方程

$$3x^2+y^2=z^2$$

的正整数解。根据方程(1)解的形式，我们猜想该方程的解的形式为

$$x = 2mn, y = |3m^2 - n^2|, z = 3m^2 + n^2。$$

我们可仿照上面的证明过程，证明这个结论是对的。

在证明过程中应用了 3 是素数的性质，因此我们可猜想，对于一般的素数 p，不定方程

$$px^2 + y^2 = z^2$$

的正整数解具有如下形式

$$x = 2mn, y = |pm^2 - n^2|, z = pm^2 + n^2。$$

经理论推证，正是这样，于是我们有如下的结论。

定理 1.3.7 不定方程

$$px^2 + y^2 = z^2 \tag{5}$$

满足 p 是奇素数，且

$$(x, y) = 1, x > 0, y > 0, z > 0, 2 \mid x \tag{6}$$

的全部正整数解可表为

$$x = 2mn, y = |pm^2 - n^2|, z = pm^2 + n^2, \tag{7}$$

其中 $p \nmid n, m > 0, n > 0, (m, n) = 1, m, n$ 中一奇一偶。 (8)

证明 设 m，n 满足(8)，将(7)代入(5)，得

$$px^2 + y^2 = p(2mn)^2 + (pm^2 - n^2)^2$$
$$= pm^2 + n^2 = z^2。$$

显然 $x>0$，$y>0$，$z>0$，$2 \mid x$。下面证明 $(x, y) = 1$。为此设 $(x, y) = d$，再证 $d = 1$ 即可。于是有 $d \mid (pm^2 - n^2), d \mid (pm^2 + n^2)$，从而推知 $d \mid 2(pm^2, n^2)$。由 $(m, n) = 1$，知 $(m^2, n^2) = 1$，所以 $(pm^2, n^2) = (p, n^2)$。

又因为 $p \nmid n$，即 $(p, n) = 1$，所以 $(p, n^2) = 1$

故 $\quad 2(pm^2, n^2) = 2$，

即 $\quad d = 2$。

所以 $\quad d = 1$ 或 2。

因为 m, n 中一奇一偶，易知 $pm^2 - n^2$ 是奇数，而 $d \mid (pm^2 - n^2)$，所以 $d = 1$。这就证明了 $(x, y) = 1$。

反之，若(5)满足(6)的任一组解 x，y，z，一定可由(7)与

(8)表示出来。由 $2|x$ 和 $(x, y)=1$ 知 $2\nmid y$。再由 (5) 知 $2\nmid z$。所以 $y+z$, $y-z$ 都是偶数。从而推知 $(y+z, z-y)\mid 2(y,z)$。另一方面，由(5)及 $(x, y)=1$ 可推知 $(y, z)=1$。
所以：$(y+z, z-y)=2$。

于是有 $$\left(\frac{y+z}{2}, z-y\right)=1, \tag{9}$$

或 $$\left(y+z, \frac{z-y}{2}\right)=1。$$

先讨论(9)式。由(5)，得

$$px^2=\frac{y+z}{2}\cdot 2(z-y)。\tag{10}$$

由 $\left(z-y, \frac{y+z}{2}\right)=1$ 及 $z-y$ 是偶数，知 $\frac{y+z}{2}$ 是奇数，于是

$$\left(\frac{y+z}{2}, 2\right)=1。$$

故推知 $\left(\frac{y+z}{2}, 2(z-y)\right)=1$。

下面分两种情况讨论。

i) 如果 $p\left|\frac{y+z}{2}\right.$，则令 $\frac{y+z}{2}=pk$，
于是有

$$x^2=k\cdot 2(z-y)。$$

由 $\left(\frac{y+z}{2}, 2(z-y)\right)=1$，推出

$$(k, 2(z-y))=1。$$

进而可推知

$$k=m^2, 2(z-y)=(2n)^2,$$

取 m，n 同号，即有

$$z+y=2pm^2,\quad z-y=2n^2,$$
$$x^2=m^2\ (2n)^2。\tag{11}$$

由 (11) 解出 x，y，z，得

$$x=2mn,\quad y=pm^2-n^2,\quad z=pm^2+n^2。$$

显然$(m,n)=1$。若不然，就会由$(m,n)>1$导出$(x,y)>1$，与假设矛盾。另外，有$p\nmid n$。若不然，会导出$p\mid(x,y)=1$。

因为$\frac{z+y}{2}=pm^2$，而$\frac{z+y}{2}$是奇数，知m是奇数。

又由$2\nmid z$，$z=pm^2+n^2$知n是偶数。

由$z+y=2pm^2$，$z-y=2n^2$可知

$$y=pm^2-n^2>0。$$

ii）如果$p\mid 2(z-y)$，令$2(z-y)=pk$（其中k为偶数）。

因为
$$px^2=\frac{y+z}{2}\cdot 2(z-y),$$

所以
$$x^2=\frac{y+z}{2}\cdot k。$$

因为$\left(\frac{y+z}{2},z-y\right)=1$，又$\frac{y+z}{2}$是奇数，推出$\left(\frac{y+z}{2},2(z-y)\right)=1$。

所以$\left(\frac{y+z}{2},pk\right)=1$，故知$\left(\frac{y+z}{2},k\right)=1$。于是推知

$$\frac{y+z}{2}=n^2,k=(2m)^2,\text{取 } m,\ n \text{ 同号。}$$

所以$x^2=4m^2n^2$，$y+z=2n^2$，$z-y=2pm^2$。

由上述三个方程可解出x，y，z为

$$x=2mn,\quad y=n^2-pm^2,\quad z=pm^2+n^2。$$

类似i）的讨论，可知m，n满足(8)式。

下面讨论(10)式。由(5)可变形为

$$px^2=2(z+y)\cdot\frac{z-y}{2}。$$

分两种情况予以讨论：

i）如果$p\mid 2(z+y)$，令$2(z+y)=pk$（k为偶数），即有

$$x^2=k\cdot\frac{z-y}{2}。$$

易知$\left(k,\frac{z-y}{2}\right)=1$，于是可推知

$$k=(2m)^2, \frac{z-y}{2}=n^2, \text{取 } m, \ n \text{ 同号。}$$

从而得到

$$z+y=2pm^2, \quad z-y=2n^2,$$

$$x^2=4m^2n^2。$$

上述三个方程联立，得

$$x=2mn, \quad y=pm^2-n^2, \quad z=pm^2+n^2。$$

显然 $(m,n)=1, m>0, n>0, p\nmid n$ 。因为$\frac{z-y}{2}$是奇数，而 $n^2=\frac{z-y}{2}$，所以 n 是奇数。又因为$(x, y)=1$，即

$$(2mn, pm^2-n^2)=1,$$

而 n 是奇数，故推知 m 必为偶数。

ii）当 $p \mid \frac{z-y}{2}$ 时，类似 i）可证。

综上所述，知不定方程(5)满足(6)的全部正整数解可写成如下形式：

$$x=2mn, y=|pm^2-n^2|, z=pm^2+n^2。$$

其中 $m>0$，$n>0$，$(m, n)=1$，m，n 中一奇一偶。(证完)

在（6）式中有 $2|x$，若 $2\nmid x$，又有怎样的结论呢？于是得到如下定理。

定理 1.3.8 不定方程(5)

$$px^2+y^2=z^2$$

满足 p 是奇素数，且

$$(x,y)=1, x>0, y>0, z>0, 2\nmid x \tag{12}$$

的全部正整数解可写成如下形式：

$$x=mn, y=\left|\frac{pm^2-n^2}{2}\right|, z=\frac{pm^2+n^2}{2}, \tag{13}$$

其中

$$m>0, \quad n>0, \quad (m, n)=1,$$

$$p\nmid n, \quad m \text{ 与 } n \text{ 均为奇数。} \tag{14}$$

证明 设 m，n，p 满足(14)式，将(13)代入(5)，经具体计算知 x，y，z 为(5)的整数解，且满足(12)式。事实上，由 $pm^2-n^2\neq0$ 知 $y>0$。显然有 $x>0$，$z>0$，$2\nmid x$。下面再证 $(x,y)=1$。为此设 $(x,y)=d$，于是有

$$d\left|\frac{pm^2-n^2}{2}\neq d\right|\frac{pm^2+n^2}{2},$$

所以 $d\mid pm^2, d\mid n^2$，因而 $d\mid(pm^2,n^2)$。又因为 $(m^2,n^2)=1$，$(p,n^2)=1$，所以 $d=1$，即 $(x,y)=1$。

反过来，设 x，y，z 是(5)满足(12)的一组解。由 $2\nmid x$，p 是奇素数，再由(5)知 y，z 中一奇一偶。

易知 $(y+z,z-y)\mid2(y,z)$，而 $(y,z)=1$，所以 $(y+z,z-y)=2$，故知

$$(y+z,z-y)=1$$

或 $$(y+z,\ z-y)=2。$$

但 $y+z$，$z-y$ 均为奇数，所以只能是

$$(y+z,z-y)=1。$$

因为 $px^2=(y+z)(z-y)$，所以分两种情况讨论：

i) 如果 $p\mid(y+z)$，令 $y+z=pm^2$，则 $z-y=n^2$，取 m，n 同号。我们有 $x^2=m^2n^2$。解下述的方程组

$$\begin{cases}y+z=pm^2,\\ z-y=n^2,\\ x^2=m^2n^2,\end{cases}$$

得

$$x=mn,\quad y=\frac{pm^2-n^2}{2},\quad z=\frac{pm^2+n^2}{2}。$$

下面证明 m，n，p 满足(14)式。显然有 $m>0$，$n>0$，又 $(y+z,z-y)=1$，即 $(pm^2,n^2)=1$，由此推知 $(m,n)=1$。又知 $2\nmid x$，故 m，n 均为奇数。易知 $p\nmid n$。

ii) 如果 $p\mid(z-y)$，令 $z-y=pm^2$，则 $y+z=n^2$，取 m，

n 同号。解下述的方程组

$$\begin{cases} z-y=pm^2, \\ y+z=n^2, \\ x^2=m^2n^2, \end{cases}$$

得 $x=mn,\quad y=\dfrac{n^2-pm^2}{2},\quad z=\dfrac{pm^2+n^2}{2}$。

仿前面讨论知 m，n，p 满足(14)式。

综合上述，知定理成立。(证完)

上述不定方程（5）中是对 x 前面的系数作了限制而求出不定方程的正整数解来。如果 x 前面的系数 p 不是奇素数，而是一般的任意数 a，怎样求不定方程的整数解呢？即求不定方程

$$ax^2+y^2=z^2 \tag{15}$$

的整数解。

解　(15) 经变形，得

$$ax^2=z^2-y^2,$$

即 $$ax^2=(z+y)(z-y)。$$

设 $a=MN$，$z+y=2Mu^2$，$z-y=2Nv^2$，则 $MNx^2=4MNu^2v^2$，即

$$x^2=4u^2v^2。$$

由此，求得

$$\begin{cases} x=\dfrac{2}{F}uv, \\ y=\dfrac{1}{F}(Mu^2-Nv^2), \\ z=\dfrac{1}{F}(Mu^2+Nv^2), \end{cases} \tag{16}$$

其中 F 为各式的最大公因数，u 与 v 是互素的整数，$MN=a$。

方程（15）中的 a，如果能写成两数的平方差的形式，即 $a=M^2-N^2$，则 x，y，z 可表示成另外的形式。设 $a=M^2-N^2$，则

$$(M^2-N^2)x^2+y^2=z^2,$$

即 $$(y+Nx)(y-Nx)=(z+Mx)(z-Mx)。$$

令
$$\begin{cases}y+Nx=u, y-Nx=g,\\ z+Mx=h, z-Mx=v,\end{cases}\tag{17}$$

则有 $$ug=hv。\tag{18}$$

由(17)，得

$$\begin{cases}y=\frac{1}{2}(u+g), x=\frac{1}{2N}(u-g),\\ z=\frac{1}{2}(h+v), x=\frac{1}{2M}(h-v),\end{cases}\tag{19}$$

其中 $h=\frac{M}{N}(u-g)+v$。 (20)

将(20)代入(18)，得

$$g=\frac{v\ (Mu+Nv)}{Nu+Mv}。$$

将上式代入（20），得

$$h=\frac{u(Mu+Nv)}{Nu+Mv}。$$

将 g，h 的表达式代入（19），得

$$x=\frac{u^2-v^2}{2(Nu+Mv)},$$

$$y=\frac{Nu^2+2Muv+Nv^2}{2(Nu+Mv)},$$

$$z=\frac{Mu^2+2Nuv+Mv^2}{2(Nu+Mv)}。$$

于是，
$$\begin{cases}x=\frac{1}{F}(u^2-v^2),\\ y=\frac{1}{F}(Nu^2+2Muv+Nv^2),\\ z=\frac{1}{F}(Mu^2+2Nuv+Mv^2),\end{cases}\tag{21}$$

其中 $M^2-N^2=a$。

上面仅讨论了一个元的系数不是 1 的情况，如果有两个元的

系数不是 1 的情况，其不定方程的整数解又怎样去求呢？即求不定方程

$$ax^2 + by^2 = z^2 \tag{22}$$

的整数解。

解 设 a 是一个非平方正整数，且方程

$$X^2 - aY^2 = b \tag{23}$$

有正整数解存在。设 X_1，Y_1 是方程（23）的一组正整数解，则（22）可写为

$$ax^2 + (X_1^2 - aY_1^2)y^2 = z^2,$$

经变形，得

$$a(x + Y_1 y)(x - Y_1 y) = (z + X_1 y)(z - X_1 y)。$$

令
$$\begin{cases} x + Y_1 y = u, & x - Y_1 y = g, \\ z + X_1 y = h, & z - X_1 y = v, \end{cases}$$

则有

$$aug = hv。$$

仿照上述的方法予以推导，求得

$$\begin{cases} x = \dfrac{1}{F}(aY_1 u^2 + 2X_1 uv + Y_1 v^2), \\ y = \dfrac{1}{F}(au^2 - v^2), \\ z = \dfrac{1}{F}(aX_1 u^2 + 2aY_1 uv + X_1 v^2), \end{cases} \tag{24}$$

其中 X_1，Y_1 是方程 $X^2 - aY^2 = b$ 的一组正整数解。

若 b 是一个非平方整数，且方程 $X^2 - bY^2 = a$ 存在正整数解，设其解为 X_1，Y_1，同样可求得

$$\begin{cases} x = \dfrac{1}{F}(bu^2 - v^2), \\ y = \dfrac{1}{F}(bY_1 u^2 + 2X_1 uv + Y_1 v^2), \\ z = \dfrac{1}{F}(bX_1 u^2 + 2Y_1 uv + X_1 v^2)。 \end{cases} \tag{25}$$

上面我们从方程中的系数由相同到不同而提出一些不定方程求解问题。如果从方程 $x^2+y^2=z^2$ 中的指数这个角度来考虑也可以提出一些新的问题出来。方程中的指数都是 2，若指数不全都是 2，又会怎样呢？比如我们可提出求不定方程

$$x^4+y^4=z^2$$

的正整数解的问题。关于这一问题，我们在后面将专门来研究。类似的问题还可以提出很多。

在一些含有各种量的数学问题中，由于量之间的相同与不相同，就形成了可能是不同类型的问题。当我们考虑了其中一种类型的问题之后，就可提出其他类型的问题。从量的相同到不同可提出新问题来，反过来，也可以从量的不同到相同提出问题来。这种由量的异同提出新问题的方法也是从数学内部提出问题的重要方法之一。

3. 由多到少，由少到多

不定方程 $x^2+y^2=z^2$ 有三个未知元，如果元数减少为两个，比如令 $z^2=1$，这时方程变成

$$x^2+y^2=1。$$

显然，该方程恒有二组非负整数解：

$$x=1,\quad y=0;$$

$$x=0,\quad y=1。$$

上述方程中的两个元的系数都是 1，如果其中一个元的系数不是 1，比如 y 前面的系数为 b，我们来研究方程的整数解，即求不定方程

$$x^2+by^2=1$$

的整数解。该方程叫做 Pell 方程。

若 x，y 是上述方程的解（为了叙述简单，以下所谈方程的解，都指整数解），则 x，$-y$；$-x$，y；$-x$，$-y$ 这三组数也都是上述方程之解，所以我们只要讨论 Pell 方程的非负解就够

了。

当 $b>0$ 时，方程只有 $x=1$，$y=0$ 这一组解，所以又只要研究 b 为负整数的情况。

若 $|b|$ 是一个完全平方数，设 $|b|=c^2$，则

$$x^2+by^2=x^2-(cy)^2=(x+cy)(x-cy)=1,$$

即 $(x+cy)|1,(x-cy)|1$，故只有 $x=1$，$cy=0$，即有一组非负整数解 $x=1$，$y=0$。

所以我们的问题又归结到只要研究 $b<0$，且 $|b|$ 不是一完全平方数的情况。为此，设 $b=-B$，有下面结论。

定理 1.3.9 设 B 是一个正整数，且不是一个完全平方数，则方程

$$x^2-By^2=1 \tag{1}$$

有无限多组整数解 x，y。

设 $x_0^2-By_0^2=1$，$x_0>0$，$y_0>0$ 是所有 $x>0$，$y>0$ 的解中使 $x+y\sqrt{B}$ 最小的那组整数解（称 x_0，y_0 为（1）的基本解），则(1)的全部解 x，y 可写成如下形式：

$$x+y\sqrt{B}=\pm(x_0+y_0\sqrt{B})^n, \tag{2}$$

其中 n 是任意整数。

为了证明这个定理，可将其分解成几个基本问题。如果基本问题解决了，那么定理也就容易证明了。这些基本问题，我们以引理形式表出。

引理 1 设 θ 是无理数，且 $q>1$ 是任意给定的整数。设 $L=x-y\theta$，则存在整数 x，y，使得

$$|L|<\frac{1}{q},0<y\leqslant q。 \tag{3}$$

证 设 y 取 0，1，2，…，q，均存在整数 x，使 $y\theta\leqslant x<y\theta+1$，即

$$0\leqslant L<1。$$

故将 (x_i, y_i)，$i=1,2,\cdots,q+1$ 代入 L 中，得

L_i，$i=1$，2，…，$q+1$，且满足 $0\leqslant L_i<1$，$i=1$，2，…，$q+1$。对于下面的 q 个半开区间

$$\left[\frac{r}{q},\frac{r+1}{q}\right),r=0,1,2,\cdots,q-1,$$

根据抽屉原则，至少有一个区间落于两个 L 的值，不妨设这两个 L 值为 $L_1=x_1-y_1\theta$，$L_2=x_2-y_2\theta$，$y_1>y_2$，故有

$$|L_1-L_2|=|x_1-x_2-(y_1-y_2)\theta|<\frac{1}{q},$$

令　$x=x_1-x_2$，$y=y_1-y_2$，且 $0<y\leqslant q$，

故（3）式成立。　　　（证完）

推论　有无穷多对整数 x，y，适合不等式

$$|x-y\theta|<\frac{1}{y}。\tag{4}$$

证　由(3)知，有整数 x_i，y_i，适合(4)。取整数 $q_i>1$，使

$$\frac{1}{q_i}<|x_i-y_i\theta|<\frac{1}{y_i}。$$

由引理 1 知，存在 x_{i+1}，y_{i+1}，适合

$$|x_{i+1}-y_{i+1}\theta|<\frac{1}{q_i}\leqslant\frac{1}{y_{i+1}}。$$

而上式中的下足码 i 可取 1，2，…，于是有

$$|x_1-y_1\theta|>|x_2-y_2\theta|>\cdots,$$

故 x_i，y_i，$i=1$，2，…是不同的整数对，即知有无穷多对整数 x_i，y_i，适合(4)式。(证完)

引理 2　设 B 不是平方数，$B>0$，则存在无穷多对整数 x，y，使得

$$|x^2-By^2|<1+2\sqrt{3}。$$

证　在(3)中取 $\theta=\sqrt{B}$，由引理 1 的推论知，存在无穷多对整数 x，$y>0$，使得下式成立

$$|x-y\theta|<\frac{1}{y}。\tag{5}$$

又有

$$|x+y\theta| = |x-y\theta+2y\theta| \leqslant |x-y\theta|+2y\theta < \frac{1}{y}+2y\sqrt{B}。 \quad (6)$$

由(5)×(6)，得

$$|x^2-y^2\theta^2| = |x^2-By^2| < \frac{1}{y^2}+2\sqrt{B} \leqslant 1+2\sqrt{B}。$$

(证完)

引理3 设 B 不是平方数，$B>0$，则存在整数 k，$0<|k|<1+2\sqrt{B}$，使得

$$x^2-By^2=k \quad (7)$$

有无穷多组整数解 x，y。

证 绝对值小于 $1+2\sqrt{B}$ 的整数只能是有限个，根据引理2知，存在整数 k，且 $|k|<1+2\sqrt{B}$，使得(7)有无穷多组整数解 x，y。又因 B 不是平方数，故 $x^2-By^2\neq 0$，即 $|k|>0$。(证完)

推论 设 B 不是平方数，$B>0$，则存在整数 k，$0<|k|<1+2\sqrt{B}$，使得(7)有无穷多组整数解 $x>0$，$y>0$。

现在对定理1.3.9进行证明。

先证明 (1) 式至少有一组解 x，$y\neq 0$。由引理3知(7)式有无穷多组解 $x>0$，$y>0$，从而其中至少有两组不同的解 $(x_1, y_1)\neq(x_2, y_2)$，其中 x_1，x_2，y_1，y_2 皆大于零，且满足如下关系：

$$x_1 = x_2(\bmod |k|), y_1 = y_2(\bmod |k|), \quad (8)$$

于是有

$$(x_1^2-By_1^2)(x_2^2-By_2^2) = (x_1x_2-By_1y_2)^2 - B(x_1y_2-x_2y_1)^2 = k^2。 \quad (9)$$

设 $x_1x_2-By_1y_2=Xk$，$x_1y_2-x_2y_1=Yk$。由 (9) 得

$$X^2-BY^2=1。$$

下面证明 X，Y 是整数。由(8)得

$$x_1x_2 - By_1y_2 \equiv x_1^2 - Dy_1^2 = k \equiv 0 (\bmod |k|),$$
$$x_1y_2 - x_2y_1 \equiv x_2y_2 - x_2y_2 = 0 (\bmod |k|),$$

故 X，Y 是整数，且 $Y \neq 0$。如若不然，由 $Y=0$，知 $x_1y_2 = x_2y_1$，即 $\frac{x_1}{x_2} = \frac{y_1}{y_2}$。设此比值为 $l>0$，则有 $x_1 = x_2 l$，$y_1 = y_2 l$。代入（7）得

$$k = l^2(x_2^2 - Dy_2^2) = l^2 k。$$

所以 $l=1$，故 $x_1 = x_2$，$y_1 = y_2$，这与 $(x_1, y_1) \neq (x_2, y_2)$ 相矛盾。由此得知 X，Y 是（1）的一组解，且 $Y \neq 0$。不失一般性，可设 $X>0$，$Y>0$。设 x_0，y_0 是（1）的基本解，则满足

$$x + y\sqrt{B} = (x_0 + y_0\sqrt{B})^n, n > 0 \tag{10}$$

的 x，y 是(1)的解。记 $E = x_0 + y_0\sqrt{B}$，$\bar{E} = x_0 - y_0\sqrt{B}$。因为对于任意给定的整数 $n>0$，有 $\bar{E} = x - y\sqrt{B}$，$x^2 - y^2B = (E\bar{E})^n = 1$，故给出(1)的一组解 $x>0$，$y>0$。又因为 $E>1$，所以不同的 n 给出的解也不相同。于是(10)给出(1)的无穷组解 $x>0$，$y>0$。反之，(1)的任一组解 $x>0$，$y>0$ 都可以写成(10)的形式。如若不然，则有 $x + y\sqrt{B} > x_0 + y_0\sqrt{B}$，必存在某个整数 n，使得

$$E^n < x + y\sqrt{B} < E^{n+1}。$$

上式两端同乘以 $\bar{E}^n$，得

$$1 < (x + y\sqrt{B})\bar{E}^n < E。$$

$(x + y\sqrt{B})\bar{E}^n$ 是属于 $u + v\sqrt{B}$ 的形式，显然 u，v 是一组解。由于

$$u + v\sqrt{B} > 1, \tag{11}$$

故

$$0 < u - v\sqrt{B} = \frac{1}{u + v\sqrt{B}} < 1。 \tag{12}$$

由（11）+（12），得

$$2u > 1 + 0 = 1,$$

故 $u>0$。

由（11）-（12），得

$$2v\sqrt{B}>1-1=0,$$

故 $v>0$。而 $u+v\sqrt{B}<E$，此与 E 的选择发生矛盾。这就证明了(1)的全体解 $x>0$，$y>0$ 可以用(10)式表示出来。

用上述的结果，(1)的全体解 $x<0$，$y<0$，可以表示为

$$x+y\sqrt{B}=-E^n,\quad n>0, \tag{13}$$

(1) 的全体解 $x<0$，$y>0$，可表为

$$x+y\sqrt{B}=-E^{-n},\quad n>0, \tag{14}$$

(1) 的全体解 $x>0$，$y<0$，可表为

$$x+y\sqrt{B}=E^{-n},\quad n>0, \tag{15}$$

(1) 的平凡解 $x=\pm1$，$y=0$，可表为

$$x+y\sqrt{B}=E^0。 \tag{16}$$

由(10)，(13)，(14)，(15)，(16) 知，(1)的全体解 x，y 可写成(2)的形式。(证完)

推论 对于任意给定的整数 $n>0$，(1) 存在无穷多组解 x，y，满足 $y\equiv0 \pmod{n}$。

定理 1.3.9 告诉我们只要求出(1)的基本解，则其全部解也就求出来了。怎样求(1)的基本解呢？这可用试验的方法。设 $y=1,2,3,\cdots$，直到 $1+By^2$ 是一个完全平方数，就可求出基本解。

例 求出 $x^2-10y^2=1$ 的全部整数解。

解 先求出基本解 x_0，y_0。

用试验的方法，设 $y=1$，2，3，…。

当 $y=1$ 时，$1+10y^2=11\neq x^2$。

当 $y=2$ 时，$1+10y^2=41\neq x^2$。

当 $y=3$ 时，$1+10y^2=91\neq x^2$。

当 $y=4$ 时，$1+10y^2=161\neq x^2$。

当 $y=5$ 时，$1+10y^2=251\neq x^2$。

当 $y=6$ 时，$1+10y^2=361=19^2$。

所以 $x^2-10y^2=1$ 的基本解为 $x_0=19$，$y_0=6$。

再写出 x，y 的全部解。

$$x+\sqrt{10}y=\pm\ (x_0+y_0\sqrt{10})^n=\pm\ (19+6\sqrt{10})^n。$$

（1）式仅含两个元，如果将元数增多，也可提出许许多多的问题，比如可提出如下的问题：求四元二次不定方程

$$x_1^2+x_2^2+x_3^2+x_4^2=1$$

的非负整数解。

显然，此方程有下列四组非负解：

$$x_1=1,\ x_2=0,\ x_3=0,\ x_4=0;$$
$$x_1=0,\ x_2=1,\ x_3=0,\ x_4=0;$$
$$x_1=0,\ x_2=0,\ x_3=1,\ x_4=0;$$
$$x_1=0,\ x_2=0,\ x_3=0,\ x_4=1。$$

一个数学问题中含有各种量，由于量的数目不一样，往往就可以形成不同类型的问题。因此，当我们解决其中一种类型的问题之后，就可从问题所含量的数量增减提出问题。对所提出的问题进行研究与解决，就可丰富数学的内容。以量的增减提出问题也是从数学内部提出新问题的方法之一。

4. 由特殊到一般，由一般到特殊

2 中的(22)式相对于不定方程 $x^2+y^2=z^2$ 是一般形式，但相对于

$$ax^2+by^2=cz^2, \tag{1}$$

又是一个当 $c=1$ 时的特殊情况。当我们考虑了（22）式的整数解之后，由特殊到一般就可提出更一般的问题：求不定方程(1)的整数解。

我们首先研究此类方程有整数解的条件，然后给出求解的方法。

显然，(1)有平凡解 $x=0$，$y=0$，$z=0$。我们只讨论(1)不全为零的解，且设 $(x,y,z)=1$。我们再假设 $abc\neq0$，a，b，c 全是无平方因子的整数，$(a,b,c)=1$。进一步我们可假定 $(a,b)=(a,c)=(b,c)=1$。如若不然，设 $(a,b)=d$，(a,c)

$= e, (b,c) = f$。由 $(a,b,c)=1$,可得$(d,e)=(d,f)=(e,f)=1$,且由 a, b, c 无平方因子和 (1) 可得 $d \mid z, e \mid y, f \mid x$,于是,令

$$a = dea_1,\quad b = dfb_1,\quad c = efc_1,\quad x = fx_1,$$
$$y = ey_1,\quad z = dz_1,$$

代入 (1),得

$$a_1 f x_1^2 + b_1 e y_1^2 = c_1 d z_1^2, \tag{2}$$

其中$(a_1 f, b_1 e) = (a_1 f, c_1 d) = (b_1 e, c_1 d) = 1$。这样求(1)的解化为求系数两两互素的方程(2)的解。不失一般性,(1)还可进一步假定 $a>0$, $b>0$, $c>0$。

有了上述条件限定之后,虽然研究的方程相对于(1)来说是特殊了,但所得到的结论很易转化到一般方程(1)上来。下面我们就给出限定条件方程的整数解的判别条件定理。

定理 1.3.10 不定方程

$$ax^2 + by^2 = cz^2 \tag{3}$$

的系数满足条件

$$a > 0, b > 0, c > 0, (a,b) = (a,c) = (b,c) = 1,$$
$$a,\ b,\ c \text{ 都无平方因子} \tag{4}$$

时,则(3)有一组不全为零的整数解 x, y, z, 且有 $(x,y,z)=1$ 的充分必要条件是

$$\left(-\frac{ab}{c}\right) = 1, \left(\frac{bc}{a}\right) = 1, \left(\frac{ac}{b}\right) = 1。\tag{5}$$

$\left(\frac{a}{p}\right)$ 叫做勒让德 (Legendre) 符号:

$$\left(\frac{a}{p}\right) = \begin{cases} 1, \text{若 } a \text{ 是模 } p \text{ 的平方剩余;} \\ -1, \text{若 } a \text{ 是模 } p \text{ 的平方非剩余。} \end{cases}$$

假设 $(a,m)=1$,如果同余式

$$x^2 \equiv a \pmod{m}$$

有解,则 a 叫做模 m 的平方剩余,否则叫做模 m 的平方非剩余。

证 设满足条件(4)的(3)有一组解 x，y，z 不全为零，且 $(x, y, z)=1$。当 $z\neq 0$ 时，分下面几种情况讨论。

i）当 $x=0$，$y=0$ 时，这种情况不可能出现。否则，因 $a>0$，$b>0$，$c>0$，必推出 $z=0$，这与 $z\neq 0$ 的假定矛盾。

ii）当 $x=0$，$y\neq 0$ 时，则 $c\mid by^2$，因为 $(c,b)=1$，故 $c\mid y^2$。现证 $c=1$。如若不然，设 $c>1$，则有素数 $p\mid c, p\mid y^2$，推出 $p^2\mid cz^2$。因 c 无平方因子，所以 $p\mid z$。这与 $(z,y)=1$ 矛盾。这就证明了 $c=1$。于是有 $\left(-\dfrac{ab}{c}\right)=1$。

iii）当 $x\neq 0$，$y=0$ 时，类似 ii）可证。

iv）当 $x\neq 0$，$y\neq 0$ 时，由 $ax^2+by^2\equiv 0 \pmod{c}$，$(aby^2, c)=1$，得

$$\left(\frac{-aby^2}{c}\right)=\left(-\frac{ab}{c}\right)=1。$$

类似地，可证得 $\left(\dfrac{bc}{a}\right)=\left(\dfrac{ac}{b}\right)=1$。必要性证完。

现证充分性。设(3)满足条件(4)，且(5)成立。

当 $a=b=c=1$，(3)显然有不全为零的解。下面讨论 a，b，c 不全为 1 的情况。由(3)取模 c，得

$$ax^2+by^2\equiv 0 \pmod{c} \tag{6}$$

因 $\left(\dfrac{-ab}{c}\right)=1$，$(a,c)=1$，故有整数 k，使得 $k^2\equiv -ab \pmod{c}$；有整数 a_1，使 $aa_1\equiv 1 \pmod{c}$。于是由(6)，得

$$ax^2+by^2-cz^2\equiv ax^2+by^2\equiv a_1a^2x^2+a_1aby^2$$
$$\equiv a_1(a^2x^2-k^2y^2)=a_1(ax-ky)(ax+ky) \pmod{c}。$$

由（3）取模 a，得

$$by^2-cz^2\equiv 0 \pmod{a}。 \tag{7}$$

因为 $\left(\dfrac{bc}{a}\right)=1$ 和 $(a,b)=1$，故存在整数 S，使 $S^2\equiv bc \pmod{a}$；有整数 b_1，使 $bb_1\equiv 1 \pmod{a}$。于是由(7)，得

$$ax^2+by^2-cz^2\equiv by^2-cz^2\equiv b_1b^2y^2-b_1bcz^2$$

$$\equiv b_1(b^2y^2 - S^2z^2) = b_1(by - Sz)(by - Sz)(\mathrm{mod}\,a)。$$

由(3)取模 b，得

$$ax^2 - cz^2 \equiv 0(\mathrm{mod}\,b)。\tag{8}$$

因为 $\left(\frac{ac}{b}\right)=1$ 和 $(a,b)=1$，故存在 t，使 $t^2\equiv ac(\mathrm{mod}\,b)$ 和 a_2 使 $aa_2\equiv 1(\mathrm{mod}\,b)$。于是由(8)，得

$$ax^2 + by^2 - cz^2 \equiv a_2(ax - tz)(ax + tz)(\mathrm{mod}\,b)。$$

我们将上面得到的几个式子，用统一符号记，即存在

$$L_i(x,y,z) = l_ix + m_iy + n_iz,$$

$$i=1,\ 2,\ 3$$

$$M_i(x,y,z) = u_ix + v_iy + w_iz,$$

使

$$ax^2 + by^2 - cz^2 \equiv L_i(x,y,z)M_i(x,y,z)(\mathrm{mod}\,a_i)$$

$$i=1,\ 2,\ 3,\ a_1=a,\ a_2=b,\ a_3=c。$$

由孙子定理①，存在整数 l，m，n 和 u，v，w，满足

$$l\equiv l_1(\mathrm{mod}\,a), l\equiv l_2(\mathrm{mod}\,b), l\equiv l_3(\mathrm{mod}\,c);$$

$$m\equiv m_1(\mathrm{mod}\,a), m\equiv m_2(\mathrm{mod}\,b), m\equiv m_3(\mathrm{mod}\,c);$$

$$n\equiv n_1(\mathrm{mod}\,a), n\equiv n_2(\mathrm{mod}\,b), n\equiv n_3(\mathrm{mod}\,c);$$

$$u\equiv u_1(\mathrm{mod}\,a), u\equiv u_2(\mathrm{mod}\,b), u\equiv u_3(\mathrm{mod}\,c);$$

$$v\equiv v_1(\mathrm{mod}\,a), v\equiv v_2(\mathrm{mod}\,b), v\equiv v_3(\mathrm{mod}\,c);$$

$$w\equiv w_1(\mathrm{mod}\,a), w\equiv w_2(\mathrm{mod}\,b), w\equiv w_3(\mathrm{mod}\,c)。$$

---

① 孙子定理　设 $n\geqslant 2$，m_1，m_2，…，m_n 是两两互素的正整数。令

$$m_1m_2\cdots m_n = M = m_1M_1 = m_2M_2 = \cdots = m_nM_n,$$

则同余组

$$\begin{cases} x\equiv c_1(\mathrm{mod}\,m_1), \\ x\equiv c_2(\mathrm{mod}\,m_2), \\ \cdots\cdots \\ x\equiv c_n(\mathrm{mod}\,m_n) \end{cases}$$

有且只有解

$$x\equiv M_1a_1c_1 + M_2a_2c_2 + \cdots + M_na_nc_n(\mathrm{mod}\,M),$$

其中　$M_ka_k\equiv 1(\mathrm{mod}\,m_k)$，$k=1,\ 2,\ \cdots,\ n$。

令 $L(x,y,z) = lx + my + nz,$

$M(x,y,z) = ux + vy + wz,$

则有

$$ax^2 + by^2 - cz^2 \equiv L(x,y,z)M(x,y,z)(\bmod abc)。$$

考虑整数组成的三元有序集

$$T = \{(x,y,z) \mid 0 \leqslant x < \sqrt{bc}, 0 \leqslant y < \sqrt{ac}, 0 \leqslant z < \sqrt{ab}\}。$$

由条件（4）知，$\sqrt{ab}$，$\sqrt{bc}$，$\sqrt{ca}$中至少有一个无理数，设为$\sqrt{ca}$。因此 x 在 T 中取 $[\sqrt{ca}+1]$ 个值，由此推出 T 中元素的个数 $\geqslant \sqrt{bc}\sqrt{ab}(1+[\sqrt{ca}]) > abc$。于是有不同的两元素 (x_1, y_1, z_1)，$(x_2, y_2, z_2) \in T$，使

$$L(x_1,y_1,z_1) \equiv L(x_2,y_2,z_2)(\bmod abc)。$$

因此，$L(x_1-x_2, y_1-y_2, z_1-z_2) \equiv 0(\bmod abc)$。

设 $|x_1-x_2| = x$，$|y_1-y_2| = y$，$|z_1-z_2| = z$，即有一组不全为零的整解，满足

$$ax^2 + by^2 - cz^2 \equiv 0\ (\bmod abc),$$

且 $0 \leqslant x < \sqrt{bc}$，$0 \leqslant y < \sqrt{ca}$，$0 \leqslant z < \sqrt{ab}$。

由上式得

$$-abc < ax^2 + by^2 - cz^2 < 2abc,$$

故得

$$ax^2 + by^2 - cz^2 = 0 \tag{9}$$

或

$$ax^2 + by^2 - cz = abc。 \tag{10}$$

由（10），得

$$ax^2 + by^2 = cz^2 + abc,$$

即 $ax^2 + by^2 = c(z^2 + ab)$。

方程两边同乘以 z^2+ab，得

$$(ax^2 + by^2)(z^2 + ab) = c(z^2 + ab)^2。$$

经变形，得

$$a(xz+by)^2+b(yz-ax)^2=c(z^2+ab)^2。$$

令 $xz+by=X$，$yz-ax=Y$，$z^2+ab=Z\neq0$。

于是有

$$aX+yY=cZ。$$

因此，由(9)或(10)给出(3)的一组不全为零的解，约去其最大公因数，就得到(3)的一组满足 $(x,y,z)=1$ 的解。(证完)

下面我们给出(1)的具体解的形式。

若方程(1)有正整数解，设为 (x_0,y_0,z_0)，则(1)可写为

$$\frac{cz_0^2-by_0^2}{x_0^2}x^2+by^2=cz^2。$$

经变形，得

$$b(x_0y+y_0x)(x_0y-y_0x)=c(x_0z+z_0x)(x_0z-z_0x)。$$

令 $$x_0y+y_0x=u,\quad x_0y-y_0x=g,$$

$$x_0z+z_0x=h,\quad x_0z-z_0x=v,$$

则有

$$bug=chv。$$

解上述方程，得

$$\begin{cases} x=\dfrac{x_0}{F}(bu^2-cv^2), \\ y=\dfrac{1}{F}(by_0u^2+2cz_0uv+cy_0v^2), \\ z=\dfrac{1}{F}(bz_0u^2+2by_0uv+cz_0v^2), \end{cases}$$

其中 $(x_0,\ y_0,\ z_0)$ 是方程 $ax^2+by^2=cz^2$ 的一组正整数解，F 是各式的最大公因数。

同理可得：

$$\begin{cases} x = \dfrac{1}{F}(ax_0u^2 + 2cz_0uv + cx_0v^2), \\ y = \dfrac{y_0}{F}(au^2 - cv^2), \\ z = \dfrac{1}{F}(az_0u^2 + 2ax_0uv + cz_0v^2), \end{cases},$$

其中 (x_0, y_0, z_0) 是(1) 的一组正整数解，F 是各式的最大公因数。

$$\begin{cases} x = \dfrac{1}{F}(ax_0u^2 + 2by_0uv - b_0x_0v^2), \\ y = \dfrac{1}{F}(ay_0u^2 - 2ay_0uv - by_0v^2), \\ z = \dfrac{z_0}{F}(au^2 + bv^2), \end{cases}$$

其中(x_0, y_0, z_0)是(1)的一组正整数解，F 是各式的最大公因数。

当我们研究了一般方程(1)之后，可再考虑它的特殊情况。特别当 $cz^2 = -1$，$a = 1$，$b = -B$（$B > 0$）时，求不定方程

$$x^2 - By^2 = -1 \tag{11}$$

的整数解。

当 $B = 1$ 时，(11)显然无整数解。当 $B = 4$ 时，(11) 亦无解。1，4 都是平方数，一般地，当 B 是平方数时，(11)都无整数解。进一步我们知，当 $B = 4k(k = 1,2,3,\cdots)$ 时，(11)亦无整数解。当 $B = 5$ 时，(11) 有整数解 $y = \pm 1$，$x = \pm 2$；当 $B = 13$ 时，有 $y = \pm 5$，$x = \pm 18$；当 $B = 17$ 时，有 $y = \pm 15$，$x = 62$。一般地，我们猜想当 $B \equiv 1 \pmod 4$ 时，(11)有整数解 x，y。经证明这一猜想是对的。

当 B 不满足上述条件时，是否 (11) 都没有整数解呢？答案是否定的。下面我们给出无整数解的条件。

定理 1.3.11 设 $B = 2p$，当 p 是一个素数，$2p = r^2 + s^2$，$r \equiv \pm 3 \pmod 8$，$s \equiv \pm 3 \pmod 8$，则方程(11)无整数解。

经过上面的探索之后，进一步我们要问：（11）如果有解，是否也是无穷多组解呢？我们可采取如下方法进行推测。

由于(11)在形式上与我们已经研究的不定方程 $x^2-By^2=1$ 只在等式左边差了一个符号"-"，因此使我们想到如果将(11)两边平方，得

$$(x^2-By^2)^2=(-1)^2,$$

经变形，得

$$(x+y\sqrt{B})^2(x-y\sqrt{B})^2=1。$$

如果设 $(x+y\sqrt{B})^2=u+v\sqrt{B}$，则由

$$u=x^2+By^2, v=2xy，得$$

$$(x-y\sqrt{B})^2=u-v\sqrt{B},$$

故 $u^2-v\sqrt{B}=1$。由此得知，若 x,y 是 $x^2-By^2=-1$ 的一组解，则 u,v 是 $x^2-By=1$ 的一组解。这样一来，(11)的解可通过 $u^2-Bv^2=1$ 求出来。由于 $u^2-Bv^2=1$ 有无穷多组解，因此我们猜想(11)如果有解，也一定有无穷多组解。经推证我们有如下结果。

定理 1.3.12 设 B 是一个正整数且不是一个完全平方数，如果(11)有整数解，且设 $x_1>0, y_1>0$ 是满足方程 $x_1^2-By_1^2=-1$ 的所有 $x>0, y>0$ 的整数解中使 $x+y\sqrt{B}$ 最小的那组解 (x_1, y_1 叫做(11)的基本解)，则(11)的全部整数解 x,y 可写成如下形式：

$$x+y\sqrt{B}=\pm(x_1+y_1\sqrt{B})^{2n+1}, \tag{12}$$

其中 n 是任意整数，且

$$x_0+y_0\sqrt{B}=(x_1+y_1\sqrt{B})^2, \tag{13}$$

其中 x_0，y_0 是方程 $x^2-By^2=1$ 的基本解。

证明 设 $(x_1+y_1\sqrt{B})^2=u+v\sqrt{B}$，则

$$(x_1-y_1\sqrt{B})^2=u-v\sqrt{B}，于是$$

$$u^2-Bv^2=(u+v\sqrt{B})(u-v\sqrt{B})$$

$$= (x_1 + y_1\sqrt{B})^2(x_1 - y_1\sqrt{B})^2$$
$$= (x_1^2 - By_1^2)^2 = (-1)^2 = 1。$$

因此 u，v 是方程

$$x^2 - By^2 = 1$$

的一组解。下面证明 (13) 给出了上述方程的基本解。如若不然，则有另一基本解 x_0，y_0，使

$$1 < x_0 + y_0\sqrt{B} < (x_1 + y_1\sqrt{B})^2,$$

将上述不等式中均乘上 $-x_1 + y_1\sqrt{B}$，即

$$-x_1 + y_1\sqrt{B} < (x_0 + y_0\sqrt{B})(-x_1 + y_1\sqrt{B})$$
$$< (x_1 + y_1\sqrt{B})^2(-x_1 + y_1\sqrt{B})$$

即
$$-x_1 + y_1\sqrt{B} < (x_0 + y_0\sqrt{B})(-x_1 + y_1\sqrt{B})$$
$$< x_1 + y_1\sqrt{B}。 \tag{14}$$

设 $(x_0 + y_0\sqrt{B})(-x_1 + y_1\sqrt{B}) = x' + y'\sqrt{B}$，则
$x' = -x_0x_1 + y_0y_1B$，$y' = x_0y_1 - y_0x_1$。因此

$$x'^2 - By'^2 = (-x_0x_1 + y_0y_1B)^2 - B(x_0y_1 - y_0x_1)^2$$
$$= x_0^2(x_1^2 - By_1^2) + y_0^2B(By_1^2 - x_1^2)$$
$$= (x_0^2 - By_0^2)(x_1^2 - By_1^2) = 1 \times (-1) = -1。$$

于是 (14) 可写为

$$0 < -x_1 + y_1\sqrt{B} < x' + y'\sqrt{B} < x_1 + y_1\sqrt{B}, y_1 \neq 0。 \tag{15}$$

因为 $x'^2 - By'^2 = -1$，故 $x' + y'\sqrt{B} \neq 1$。

如果 $1 < x' + y'\sqrt{B}$，则有

$$1 < x' + y'\sqrt{B} < x_1 + y_1\sqrt{B}。$$

而
$$x' + y'\sqrt{B} = \frac{1}{-x' + y'\sqrt{B}},$$
所以有

$$0 < -x' + y'\sqrt{B} < 1。$$

由
$$1 + 0 < x' + y'\sqrt{B} + (-x' + y'\sqrt{B}),$$

即 $1<2y'\sqrt{B}$,

知 $y'>0$。

由 $x'+y'\sqrt{B}-(-x'+y'\sqrt{B})>1-1$,

即 $2x'>0$,

知 $x'>0$。此与 $x_1^0+y_1\sqrt{B}$的定义矛盾。

如果 $x'+y'\sqrt{B}<1$,则由(15)知，有

$$1<-x'+y'\sqrt{B}<x_1+y_1\sqrt{B},$$

和 $0<x'+y'\sqrt{B}<1$。

由 $-x'+y'\sqrt{B}+x'+y'\sqrt{B}>1+0$,

知 $y'>0$;由 $-x'+y'\sqrt{B}-(x'+y'\sqrt{B})>1-1$,

知 $x'<0$。而 $(-x')^2-By'^2=-1$,

这与 $x_1+y_1\sqrt{B}$的选择发生矛盾。

由以上证明，得知(13)式成立。

对于任意整数 n,(12)式给出的 x,y,显然是(11)式的解。反过来,设 x,y 是(11)的任一组解,令

$$(x+y\sqrt{B})(-x_1+y_1\sqrt{B})=x'+y'\sqrt{B}。\qquad(16)$$

则由 $x'=-x_1x+y_1yB$，$y'=-x_1y+y_1x$，知

$$(x-y\sqrt{B})(-x_1-y_1\sqrt{B})=x'-y'\sqrt{B}。$$

于是 $(x'+y'\sqrt{B})(x'-y'\sqrt{B})=(x+y\sqrt{B})(-x_1+y_1\sqrt{B})\cdot(x-y\sqrt{B})(-x_1-y_1\sqrt{B})$,

即 $x'^2-By'^2=1$。

故知 x'，y'是不定方程

$$x^2-By^2=1$$

的一组解。于是由定理1.3.9知，x'，y'可写成如下形式:

$$x'+y'\sqrt{B}=\pm(x_0+y_0\sqrt{B})^n,$$

其中 n 为整数。(17)

由(16),(17)知,

$$(x+y\sqrt{B})(-x_1+y_1\sqrt{B})=\pm(x_0+y_0\sqrt{B})^n。\tag{18}$$

再由(13),(18)知

$$(x+y\sqrt{B})(-x_1+y_1\sqrt{B})=\pm(x_1+y_1\sqrt{B})^{2n},$$

即
$$x+y\sqrt{B}=\pm(x_1+y_1\sqrt{B})^{2n+1},$$

其中 n 为任意整数。（证完）

上述的方程 $x^2-By^2=1$，$x^2-By^2=-1$，均是特殊的二元二次方程。我们解决特殊方程 $x^2-By^2=-1$，通过问题转换，变到已解决的特殊二元二次方程 $x^2-By^2=1$ 上来。通过“特殊”解决“特殊”，这是我们解决不定方程 $x^2-By^2=-1$ 的一个基本思想方法。

当解决这两个特殊的二元二次不定方程之后，很自然地，由特殊到一般，提出一般的二元二次不定方程的整数解问题。特殊的二元二次不定方程，我们已经知道，如果它有解，一定有无穷多组解。一般寓于特殊，可猜想：对于一般的二元二次不定方程，如果它有解，一定也有无穷多组解。经过证明，这一猜想是正确的。怎样证明这一结果呢？把一般归结到特殊，通过特殊解决一般。这是解决一般的二元二次不定方程有无穷多组解的基本思想方法，下面就来叙述这一结果。

定理 1.3.13 如果满足如下条件：

i）$D=b^2-4ac>0$，D 不是一个平方数，

ii）$\Delta=4acf+bde-ae^2-cd^2-fb^2\neq 0$，

iii）不定方程

$$f(x,y)=ax^2+bxy+cy^2+dx+ey+f=0,\tag{19}$$

有一组整数解 x_0，y_0，

则(19)有无穷多组整数解，

证明 (19) $\times D^2$，得

$$aD^2x^2+bD^2xy_1+cD^2y^2+dD^2x+eD^2y+fD^2=0。\tag{20}$$

令 $Dx = x' + 2cd - be, Dy = y' + 2ae - bd$。 (21)

代入 (20) 式，得

$$a(x' + 2cd - be)^2 + b(x' + 2cd - be)(y' + 2ae - bd) + c(y' + 2ae - bd)^2 + dD(x' + 2cd - be) + eD(y' + 2ae - bd) + fD^2 = 0$$

即 $ax'^2 + bx'y' - cy'^2 - D\Delta = 0$。 (22)

由于 D 不是一平方数，所以 $a \neq 0$，不失一般性，可设 $a > 0$。(22) $\times 4a$，得

$$(2ax' + by')^2 - Dy'^2 = 4aD\Delta。 \quad (23)$$

令 $X = 2ax' + by', Y = y'$，再由(21)，得

$$X = 2aDx + bDy + dD,$$

$$Y = Dy - 2ae + bd。$$

又令 $p = 2aD$，$q = bD$，$r = dD$，$s = D$，$t = -2ae + bd$，则

$$X = px + qy + r,$$
$$Y = sy + t。 \quad (24)$$

再令 $4aD\Delta = M \neq 0$，代入(23)，得

$$X^2 - DY^2 = M。 \quad (25)$$

由于 (19) 有一组解 x_0, y_0，代入(24)，得

$$px_0 + qy_0 + r = X_0,$$
$$sy_0 + t = Y_0。 \quad (26)$$

X_0, Y_0 是(25) 的一组解。

因为 $ps > 0$，由定理 1.3.9 的推论知，对于 ps，存在无穷多个 $x^2 - Dy^2 = 1$ 的解 $u + v\sqrt{D}$，使得 $u \equiv 0 \pmod{ps}$。故有 $u^2 \equiv 1 \pmod{ps}$，且可进一步假定 $u \equiv 1 \pmod{ps}$。因为，如果 $u \not\equiv 1 \pmod{ps}$，则可取

$$u_1 + v_1\sqrt{D} = (u + v\sqrt{D})^2 = u^2 + v^2D + 2uv\sqrt{D},$$

于是有

$$u_1 = u^2 + v^2D \equiv 1 \pmod{ps},$$

$$v_1 = 2uv \equiv 0 \pmod{ps}。$$

故由 $x^2 - Dy^2 = 1$ 的无穷多组解 $u + v\sqrt{D}$，得出(25) 无穷多组解

$$\begin{aligned} &X + Y\sqrt{D} = (X_0 + Y_0\sqrt{D})(u + v\sqrt{D}), \\ &u \equiv 1 \pmod{ps},\ v \equiv 0 \pmod{ps}, \\ &X = X_0 u + Y_0 vD,\ Y = X_0 v + Y_0 u。\end{aligned} \tag{27}$$

由 (24) 和(27)，得

$$\begin{aligned} &X_0 u + Y_0 vD = px + qy + r, \\ &X_0 v + Y_0 u = sy + t。\end{aligned} \tag{28}$$

对 (28) 取模 ps，得

$$X_0 \equiv px + qy + r \pmod{ps},$$

$$Y_0 \equiv sy + t \pmod{ps}。$$

把 (26) 代入上式，得

$$p(x - x_0) + q(y - y_0) \equiv 0 \pmod{ps},$$

$$s(y - y_0) \equiv 0 \pmod{ps}。$$

于是，得

$$x = x_0 + sm_1 - qm,$$

$$y = y_0 + mp。$$

这就是说，我们从 (25) 的无穷多个整数解(27)，通过(28) 得出 (19) 的无穷多个整数解 x, y。(证完)

从上面的证明过程我们可以看出，要证明一般的 (19) 式，通过等式变形得(20) 式，再通过(21) 的变量替换，归结到它的特殊的(22) 式。继续等式变形，变量替换，又归结到它的更特殊的(25) 式。进一步归结到它的更特殊的 Pell 方程。而该方程我们已经解决了。从而一般型的不定方程的整数解的问题也就解决了。这是我们解决一般型问题的常用思想方法。

方程 $x^2 + y^2 = z^2$ 是一个二次方程，一般地，n 次方程是否有整数解？这个问题，我们将在下一节详细谈。

方程 $x^2 + y^2 = z^2$，经变形得

$$x^2+y^2-z^2=0。$$

方程 $x^2-By^2=1$，经变形，得

$$x^2-By^2-1=0。$$

方程 $x^4+y^4=z^2$，经变形，得

$$x^4+y^4-z^2=0。$$

上面等式左边的表达式都是一些特殊的整系数多项式。更一般地，设 $f(x_1,x_2,\cdots,x_n)$ 是任给的具有整系数的多项式，考虑不定方程 $f(x_1,x_2,\cdots,x_n)=0$ 的整数解的问题，就是德国数学家希尔伯特(Hilbert, D. 1862—1948) 于 1900 年提出的 23 个著名问题中的第 10 题，具体表述如下：

$f(x_1,x_2,\cdots,x_n)$ 是任给的具有整系数的多项式，试设计一种方法，根据这种方法可以通过有限步运算来判别不定方程

$$f(x_1,x_2,\cdots,x_n)=0$$

是否有有理整数解。

利用初等数论和数理逻辑的方法，给出这个问题的否定回答。

但是对于某些特殊类型的不定方程，存在一个有有限步运算的方法，来决定这些方程是否有有理整数解。

对于一个较大范围不成立的结论，我们可以缩小范围，在一个小的范围再探索结论是否成立。这也是我们思考问题的一种常规模式。

由特殊到一般，由一般到特殊，这是人类认识事物的两个基本认识过程。人们对数学的认识也遵循着这一规律。人们把这种基本认识过程用来发现数学问题，提出数学猜想，解决数学猜想，就形成了数学中的两个重要的思维原则：由特殊到一般与由一般到特殊。在一定条件下，特殊可以转化为一般，一般也可以转化为特殊。利用这种转化，就形成了数学中的两个基本方法：一般化和特殊化。这是提出数学猜想、解决数学猜想的两个重要的方法，也是数学中最常用的方法之一。

5. 由形到数，由数到形

不定方程 $x^2+y^2=z^2$ 的正整数解的问题是从几何问题勾股定理提出的，这是由“形”到“数”提出的问题。我们也可以由“数”到“形”提出问题。

我们对方程

$$x^2+y^2=z^2$$

进行变形，方程两边同除以 z^2，得

$$\left(\frac{x}{z}\right)^2+\left(\frac{y}{z}\right)^2=1。$$

令 $u=\frac{x}{z}$，$v=\frac{y}{z}$，则在（二）所讨论的问题，就变成

求圆周

$$u^2+v^2=1$$

上的有理点（其坐标为有理数的点称为有理点）。因此(二)中的定理 1.2.1 就等价于：

单位圆上有无穷多个有理点

$$u=\frac{2mn}{m^2+n^2}, v=\frac{m^2-n^2}{m^2+n^2}。$$

我们知道圆是二次曲线的特殊情况，既然圆周上有无穷多个有理点，那么我们可提出如下的猜想：

任意二次曲线上有无穷多个有理点。

这个猜想是不成立的。为了否定这一猜想，我们只要举一反例即可。例如，双曲线

$$u^2-3v^2=2 \tag{1}$$

上并没有有理点。如若不然，设

$$u=\frac{x}{z}, v=\frac{y}{z}, (x, y, z)=1,$$

则 (1) 变形为求

$$x^2-3y^2=2z^2 \tag{2}$$

之整数解的问题。

取 3 为模，则

$$x^2 \equiv 2z^2 (\mathrm{mod} 3)。$$

由此可得 $3 \mid x, 3 \mid z$。由(2)知 $3 \mid y$，这与 $(x, y, z) = 1$ 矛盾。故(2)式没有整数解，也即(1)没有有理点。

如果再加些限制条件，则有下面结果。

定理 1.3.14 在非直线的有理系数的二次曲线上如有一有理点，则有无穷个有理点。

证 不妨假定二次曲线所经过的有理点为坐标原点，不然只要作一次平移即可，所给的二次曲线可表为

$$f_2(u, v) + f_1(u, v) = 0$$

其中 $f_2(u, v)$ 为关于 u, v 的二次齐次式，$f_1(u, v)$ 为关于 u, v 的一次齐次式。$f_2(u, v)$，$f_1(u, v)$ 均不恒等于零，这是因为若 $f_2(u, v) = 0$，则原曲线为一直线，若 $f_1(u, v) = 0$，则原二次曲线为两条直线，这与题设不符。

令 $v = \omega u$，则

$$uf_2(1, \omega) + f_1(1, \omega) = 0,$$

就得

$$u = -\frac{f_1(1, \omega)}{f_2(1, \omega)}, v = -\frac{\omega f_1(1, \omega)}{f_2(1, \omega)}。$$

故有无穷多个有理点。 （证完）

定理 1.3.15 设 A, B, C 为不全为零的有理数，若 $B^2 - 4AC$ 为一平方数，则二次曲线

$$Au^2 + Buv + Cv^2 + Du + Ev + F = 0 \quad (3)$$

上有无穷个有理点。

证明 令 $B^2 - 4AC = \Delta^2$，则

$$Au^2 + Buv + cv^2 = A\left[\left(u + \frac{B}{2A}v\right)^2 + \left(\frac{C}{A} - \frac{B^2}{4A^2}\right)v^2\right]$$

$$= A\left(u + \frac{B}{2A}v - \frac{\Delta}{2A}v\right)\left(u + \frac{B}{2A}v + \frac{\Delta}{2A}v\right)。$$

若 $\Delta \neq 0$，令

$$u' = u + \frac{B + \Delta}{2A}v,\ v' = u - \frac{-B + \Delta}{2A}v,$$

解出 u，v，代入（3）式，得

$$Au'v' + D'u' + E'v' + F' = 0。$$

解出 u'，得

$$u' = -\frac{E'v' + F'}{Av' + D'}。$$

显然（3）有无穷个有理点。

若 $\Delta = 0$，令

$$u' = u + \frac{B}{2A}v,\ v' = -v,$$

则得

$$Au'^2 + D'u' + E'v' + F' = 0。$$

当 $E' \neq 0$ 时，则

$$v' = -\frac{Au'^2 + D'u' + F'}{E'}。$$

故有无穷多个有理点。

若 $E' = 0$，则原曲线不是二次曲线，故不符合题设。从而定理得证。（证完）

如果把勾股定理作为欧氏空间来考虑，那么勾股定理只是作为 n 维欧氏空间中当 $n = 2$ 时的特殊关系式。在三维欧氏空间中，直棱锥的斜面面积记为 S，三个直角面面积记为 S_1，S_2，S_3，则有等式 $S^2 = S_1^2 + S_2^2 + S_3^2$。一般地，在 n 维欧氏空间中，有类似于三维欧氏空间中的直棱锥。由一点 O 出发的两两正交的 n 个向量 $\overrightarrow{OA_1}$，$\overrightarrow{OA_2}$，…，$\overrightarrow{OA_n}$ 得到图形 $OA_1A_2\cdots A_n$ 叫做 n 维直 n 棱锥。在其各个面中，面 $A_1A_2\cdots A_n$ 叫做斜面，并用 S 表示其面积，其余的各面叫做直角面，把不含 OA_i 棱的直角面记作 S_i，则有

$$S^2 = \sum_{i=1}^{n} S_i^2。$$

我们用格拉斯曼代数证明这一等式。

证 记 $\overrightarrow{OA_i} = a_i \ | a_i | = a_i, i = 1,2,\cdots,n$，当 $i \neq j$ 时，有 $(a_i, a_j) = 0$。记 $\lambda_i = a_1 \Lambda a_2 \Lambda \cdots \Lambda a_n$，则有 $(\lambda_i, \lambda_i) = [(n-1)!\, S_i]^2$。记 $\lambda = (a_2 - a_1)\Lambda(a_3 - a_1)\Lambda\cdots\Lambda(a_n - a_1)$，则有 $(\lambda, \lambda) = [(n-1)!\, S]^2$。

由于

$$(\lambda_i, \lambda_i) = \begin{vmatrix} (a_1, a_1)(a_1, a_2)\cdots(a_1, a_{i-1})(a_1, a_{i+1})\cdots(a_1, a_n) \\ (a_2, a_1)(a_2, a_2)\cdots(a_2, a_{i-1})(a_2, a_{i+1})\cdots(a_2, a_n) \\ \cdots\cdots \quad \cdots\cdots \quad \cdots \quad \cdots\cdots \quad \cdots\cdots \quad \cdots \quad \cdots\cdots \\ (a_{i-1}, a_1)(a_{i-1}, a_2)\cdots(a_{i-1}, a_{i-1})(a_{i-1}, a_{i+1})\cdots(a_{i-1}, a_n) \\ (a_{i+1}, a_1)(a_{i+1}, a_2)\cdots(a_{i+1}, a_{i-1})(a_{i+1}, a_{i+1})\cdots(a_{i+1}, a_n) \\ \cdots\cdots \quad \cdots\cdots \quad \cdots \quad \cdots\cdots \quad \cdots\cdots \quad \cdots \\ (a_n, a_1)(a_n, a_2)\cdots(a_n, a_{i-1})(a_n, a_{i+1})\cdots(a_n, a_n) \end{vmatrix}$$

$$= \begin{vmatrix} a_1^2 & & & & & & 0 \\ & a_2^2 & & & & & \\ & & \ddots & & & & \\ & & & a_{i-1}^2 & & & \\ & & & & a_{i+1}^2 & & \\ & & & & & \ddots & \\ 0 & & & & & & a_n^2 \end{vmatrix} = (a_1^2 a_2^2 \cdots a_n^2) \cdot \frac{1}{a_i^2},$$

$i = 1,2,\cdots,n$。

$(\lambda, \lambda) =$

$$\begin{vmatrix} (a_2 - a_1, a_2 - a_1)(a_2 - a_1, a_3 - a_1)\cdots(a_2 - a_1, a_n - a_1) \\ (a_3 - a_1, a_2 - a_1)(a_3 - a_1, a_3 - a_1)\cdots(a_3 - a_1, a_n - a_1) \\ \cdots\cdots \quad \cdots\cdots \quad \cdots\cdots \\ (a_n - a_1, a_2 - a_1)(a_n - a_1)(a_3 - a_1)\cdots(a_n - a_1, a_n - a_1) \end{vmatrix}$$

$$
= \begin{vmatrix} a_2^2 + a_1^2 & a_1^2 & \cdots & a_1^2 \\ a_1^2 & a_3^2 + a_1^2 & \cdots & a_1^2 \\ \cdots & \cdots & \cdots & \\ a_1^2 & a_1^2 & \cdots & a_n^2 + a_1^2 \end{vmatrix}
$$

$$
= a_1^2 a_2^2 \cdots a_n^2 \left(\frac{1}{a_1^2} + \frac{1}{a_2^2} + \cdots + \frac{1}{a_n^2} \right)。
$$

所以 $(\lambda, \lambda) = \sum_{i=1}^{n} (\lambda_i, \lambda_i)$，从而有

$$
S^2 = \sum_{i=1}^{n} S_i^2。(证完)
$$

〔注〕记 R 是实数域，a，b，c，…是 R 中的元素。L 是R 上 n 维向量空间，α，β，γ，…是 L 中的元素。

令 $\Lambda^0 L = R$，$\Lambda^1 L = L_0$

现由 α，$\beta \in L$ 及 $a \in R$，作新元素，记为 a $(\alpha \Lambda \beta)$，且满足下列条件：

i) $(a_1\alpha_1 + a_2\alpha_2)\Lambda\beta = a_1(\alpha_1\Lambda\beta) + a_2(\alpha_2\Lambda\beta)$,

ii) $\alpha\Lambda\beta = -\beta\Lambda\alpha$。

形如 $a_i(\alpha_i\Lambda\beta_i)$ 的和 $\Sigma a_i(\alpha_i\Lambda\beta_i)$ 的全体记为 $\Lambda^2 L$。在 $\Lambda^2 L$ 中，定义加法：$\Sigma a_i(\alpha_i\Lambda\beta_i)$ 与 $\Sigma b_j(\alpha_j, \beta_j)$ 之和为 $\Sigma a_i(\alpha_i\Lambda\beta_i) + \Sigma b_j(\alpha_j, \beta_j)$；定义乘法：$a\Sigma a_i(\alpha_i, \beta_i) = \Sigma a a_i(\alpha_i, \beta_i)$。

一般地，对于 $2 \leqslant k \leqslant n$，形如 Σa $(\alpha_1\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k)$ 的全体记为 $\Lambda^k L$，且满足下列条件：

i) $(a\alpha + b\beta)\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k$

$= a(\alpha\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k) + b(\beta\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k)$;

ii) 若 $i \neq j$，$\alpha_i = \alpha_j$，则 $\alpha_1\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k = 0$;

iii) 若交换任何两个 α_i，则 $\alpha_1\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k$ 改变符号。

与 $\Lambda^2 L$ 类似，定义加法与数乘，易验证 $\Lambda^k L$ 是一个向量空间，叫做 k 一向量空间，其中的元素叫做 k 一向量。

当 $k > n$ 时，$\Lambda^k L = 0$,

现令 $G = \Lambda^0 L \cup \Lambda^1 L \cup \Lambda^2 L \cup \cdots$，那么 G 是实数域上的向量空间。

在 G 中定义乘法 Λ：

$(\alpha_1\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k)$ Λ $(Q\beta_1\Lambda\beta_2\Lambda\cdots\Lambda\beta_p)$

$= \alpha_1\Lambda\alpha_2\Lambda\cdots\Lambda\alpha_k\Lambda\beta_1\Lambda\beta_2\Lambda\cdots\Lambda\beta_p$。

于是得到一个实数域上的代数，称为格拉斯曼代数。

由上述可知，关系式 $a^2 + b^2 = c^2$ 是从二维空间中的直角三

角形概括出来的，我们从它的几何来源这一角度扩展开来，很自然地，猜想在三维空间中的直角棱锥也有类似关系式：$S^2=S_1^2+S_2^2+S_3^2$，进一步猜想：在 n 维空间中的 n 直棱锥有关系式 $S^2=\sum_{i=1}^{n}S_i^2$。要证明这一般关系式，我们不能像在二维空间、三维空间那样，利用几何方法，而是采用代数方法。当我们把形的关系式 $a^2+b^2=c^2$ 转化到数的关系式 $x^2+y^2=z^2$，从不定方程的角度出发，又提出了许许多多的问题与猜想。由此看来，数学问题的提出与解决，有的从“形”的角度较容易，也有的则从“数”的角度较方便。数与形在一定的条件可以转化，我们利用数与形这种相互转化，可以提出猜想，解决猜想。

6. 从数的性质提出问题

在直角三角形中取特定的边 3，4，5，则这三个连续的整数有等式

$$3^2+4^2=5^2。\tag{1}$$

从整数的连续性这一角度扩展开来，我们可推知有如下等式。

（1）中的等号左边是两项，右边是一项，如果在等号两边各增加一项，我们可找到五个连续整数，有

$$10^2+11^2+12^2=13^2+14^2。\tag{2}$$

在（2）中等号两边再各增加一项，可找到七个连续整数，有

$$21^2+22^2+23^2+24^2=25^2+26^2+27^2。\tag{3}$$

我们将(1)、(2)、(3) 改写为

$$(2\times1^2+1)^2+(2\times1^2+1+1)^2$$
$$=(2\times1^2+2\times1+1)^2;$$
$$(2\times2^2+2)^2+(2\times2^2+2+1)^2+(2\times2^2+2$$
$$+2)^2=(2\times2^2+2\times2+1)^2+(2\times2^2+2\times2+2)^2;$$
$$(2\times3^2+3)^2+(2\times3^2+3+1)^2+(2\times3^2+3$$

$$+2)^2+(2\times 3^2+3+3)^2=(2\times 3^2+2\times 3+1)^2$$
$$+(2\times 3^2+2\times 3+2)^2+(2\times 3^2+2\times 3+3)^2。$$

一般地，我们猜想，有

$$(2n^2+n)^2+(2n^2+n+1)^2+\cdots+(2n^2+2n)^2$$
$$=(2n^2+2n+1)^2+\cdots+(2n^2+3n)^2。\quad (4)$$

要证明这一猜想成立，只要证明下面等式

$$(2n^2+n)^2=[(2n^2+2n+1)^2-(2n^2+n+1)^2]$$
$$+[(2n^2+2n+2)^2-(2n^2+1n+2)^2]+\cdots$$
$$+[(2n^2+3n)^2-(2n^2+2n)^2]$$

即可。

而上面等式右边 $=n\ [\ (2n^2+2n+1)\ +\ (2n^2+n+1)]$

$$+n\ [\ (2n^2+2n+2)\ +\ (2n^2+n+2)]\ +\cdots$$
$$+n\ [\ (2n^2+3n)\ +\ (2n^2+2n)]$$
$$=n\ [4n^3+2n^2+n^2+2\ (1+2+\cdots+n)]$$
$$=n\left[4n^3+3n^2+2\times\frac{(1+n)\ n}{2}\right]=n\ [4n^3+3n^2+n^2$$

$+n]$

$$=n\ [4n^3+4n^2+n]\ =\ (2n^2+n)^2。$$

所以　右边 = 左边。从而猜想成立。

我们在(1)中，等号左边与右边的项数同时变动，而其乘幂的次数保持不变，由连续整数这一特点，而提出猜想(4)。如果(1)中等号右边的项不变，只变动等式左边的项，且乘幂的次数也变动，这时可推出如下的猜想。

如果只在(1)式中等号左边增加一项，而等号右边保持一项，且其次数也均升高一次，我们可找到四个连续整数，有

$$3^3+4^3+5^3=6^3。$$

一般地，我们可考虑不定方程

$$x^n+\ (x+1)^n+\cdots+\ (x+h)^n=\ (x+h+1)^n \quad (5)$$

的正整数解的问题。

当 $n>3$ 时，还没有找到方程(5)有正整数解。因此可提出

如下猜想：

不定方程(5)，除 $n=3$，$h=2$，$x=3$；$n=2$，$h=1$，$x=3$；$n=h=x=1$ 外，无正整数解。

1900 年，爱斯考特 (Escott) 证明了方程 (5) 在 $2\leqslant n\leqslant 5$ 时，除了以下情况

$$3^2+4^2=5^2,$$

$$3^3+4^3+5^3=6^3$$

外，无其他正整数解。

柯召、孙琦于 1962 年证明了 $6\leqslant n\leqslant 33$ 时，(5)无正整数解。1978 年柯召等人又证明方程(5)在 n 是奇数时，除了 $n=3$，$h=2$，$x=3$ 和 $n=h=x=1$ 外，无其它正整数解。还证明了在 $2^\beta | n$，$h\not\equiv 1$，$2\pmod{2^{\beta+3}}$，$\beta>0$ 或 $n=2^s$，$3\cdot 2^s$，$s>1$ 时，(5)均无正整数解。

在 (4) 中设 $2n^2+2n=x$，则 (4) 可改写为

$$\sum_{i=0}^{n}(x-i)^2=\sum_{i=1}^{n}(x+i)^2,$$

一般地，当 n 与 x 不一定相关，且指数也不一定是 2，这时可提出求不定方程

$$\sum_{i=0}^{h}(x-i)^m=\sum_{i=1}^{h}(x+i)^m \tag{6}$$

的正整数解的问题。

当 $m>2$ 时，找不到(6)有正整数解，于是可提出如下猜想：

当 $m>2$ 时，(6)无正整数解。

柯召于 1963 年给出这一猜想的证明。

在不定方程 $x^2-By^2=1$ 中，如果令 $B=1$，且其元的次数不一定是相同的，就可提出如下的不定方程

$$x^m-y^n=1,m>1,n>1 \tag{7}$$

的正整数解问题。x^m 与 y^n 这是两个连续的正整数，当 $x=3$，$m=2$，$y=2$，$n=3$ 时，分别有 $3^2=9$，$2^3=8$，于是有 $3^2-2^3=$

1，除此以外，再也找不到(7)式的正整数解。于是可提出如下猜想：

除开

$$m=2, x=3, y=2, n=3$$

外，(7)没有其他的正整数解。或者

不定方程

$$x^p-y^q=1, p, q \text{ 是素数,} \tag{8}$$

除开

$$p=2, x=3, y=2, q=3$$

外，没有其他的正整数解。

这个猜想换一个说法就是1842年卡塔兰（Catalan）提出的猜想：除开 $8=2^3$，$9=3^2$ 外，没有两个连续数都是正整数的乘幂。

在(8)中，当 $q=2$ 时，卡塔兰猜想成立，即有

定理 1.3.16 不定方程

$$y^2+1=x^p, p \text{ 是一个奇素数} \tag{9}$$

没有 $x>0$，$y>0$ 的整数解。

在(8)中，当 $p=2$ 时，即不定方程

$$x^2-1=y^q, q \text{ 是奇素数} \tag{10}$$

的正整数解问题，欧拉（Euler，Léonard，瑞士人，1707—1783）证明了 $q=3$ 的情况，后来的数学家相继对于 $q>3$ 的一些特殊情况作了不少工作，最后柯召在1962年彻底解决了这一问题。他得到了如下定理。

定理 1.3.17 当 $q>3$ 时，不定方程(10)无正整数解。[①]

对于（8）的正整数解的问题还可以把它转化成其他形式进行研究，为此，有如下结果。

---

① 柯召，关于方程 $x^2=y^n+1$，$xy\neq0$，四川大学学报（自然科学版），(1962)，1—6。

定理 1.3.18 不定方程

$x^p+1=y^q$，q 是素数，p 是奇素数，有正整数解的充分必要条件是

$$x+1=p^{sq-1}y_1{}^q,\ \frac{x^p+1}{x+1}=py_2^q,\ y=p^sy_1y_2,$$
$$(y_1,y_2)=1,\ p\mid y_1y_2,$$
$$y-1=q^{tp-1}x_1^p,\ \frac{y^q-1}{y-1}=qx_2^p,\ x=q^tx_1x_2,$$
$$(x_1,x_2)=1,\ q\nmid x_1x_2,$$

其中 s，t，x_1，x_2，y_1，y_2 均是正整数。

卡塔兰猜想是考虑两个连续数的正整数的乘幂，很自然地，我们可以考虑三个、四个连续正整数的乘幂问题。因为四个连续数中总有一个数呈现 $4t+2$ 型，这个数显然不能成为任何正整数的方幂。是否有三个连续数都是正整数的乘幂？这个问题叫做弱型卡塔兰猜想。因为，如果卡塔兰猜想成立，可以推出不存在三个连续数都是正整数的乘幂。关于弱型卡塔兰问题，直到 60 年代初，才由柯召和 Cassels 分别独立解决。答案是否定的，即有如下定理。

定理 1.3.19 三个正整数的乘幂不可能成为连续数。

该定理可由定理 1.3.17，定理 1.3.15 证出。

上面是根据连续整数的性质和规律提出猜想的，还可以根据数的其他性质和规律提出猜想，这里就不再赘述了。

7. 由类比提出问题

我们已经考虑了不定方程

$$x^2+y^2=z^2$$

的正整数解的问题。通过类比，由加法使我们联想到乘法，即可提出不定方程

$$x^2\times y^2=z^2$$

的正整数解的问题。而该方程显然有无穷多个正整数解。在此基础上，将指数由常量改成变量，比如可提出不定方程

$$x^x y^y = z^z \tag{1}$$

的正整数解问题。

易见，当 $x=y=z=1$；$x=1$，$y=z$；$y=1$，$x=z$ 时适合于 (1)，当 $x>1$，$y>1$，$z>1$ 就很难找到 (1) 的正整数解。于是美藉匈牙利数学家艾道斯（Erdös）在 1938 年就提出如下猜想：

没有一组正整数 $x(>1)$，$y(>1)$，$z(>1)$，适合方程(1)。

过了两年，我国著名数学家柯召证明了(1)有无穷多个解，从而否定了这一猜想。这无穷多个每个变量都大于 1 的正整数的通式为

$$x = 2^{2n+1(2n-n-1)+2n}(2^n-1)^{2(2n-1)},$$
$$y = 2^{2n+1(2n-n-1)}(2^n-1)^{2(2n-1)+2},$$
$$z = 2^{2n+1(2n-n-1)+n+1}(2^n-1)^{2(2n-1)+1},$$

其中 $n=2$，3，4，…。

如果再将问题的范围缩小一下，加上一个限制条件 $(x, y)=1$，那么猜想是成立的，即柯召证明了

定理 1.3.20 当 $(x, y)=1$ 时，(1)式无 $x>1$，$y>1$，$z>1$ 的整数解。

1958 年，施秦策（Schinzel）给出方程(1)有整数解的一个充分条件，即

定理 1.3.21 方程 (1) 有整数解，则 x 的每一个素因数整除 y，或 y 的每一个素因数整除 x。

自 1938 年以来，将近半个世纪了，至今尚没有一个找到(1)除了 $x=y=z=1$ 以外的正奇数解，于是可提出如下的猜想：

没有正奇数 $x(>1)$，$y(>1)$，$z(>1)$，适合于(1)。

方程 (1) 左边有两个未知元，如果有 k 个未知元又会怎样

呢？1964年，柯召[①]等人把(1)

推广为

$$x_1^{x_1}x_2^{x_2}\cdots x_k^{x_k}=z^z,k\geqslant 2,\tag{2}$$

他证明了

定理 1.3.22 不定方程(2)有无穷多组 $x_1(>1)$，$x_2(>1)\cdots$ $x_k(>1)$，$z(>1)$的正整数解，其通式为

$$x_1=k^{k^n(k^{n+1}-2n-k)+2n}(k^n-1)^{2(k^n-1)},$$

$$x_2=k^{k^n(k^{n+1}-2n-k)}(k^n-1)^{2(k^n-1)+2},$$

$$x_3=\cdots x_k=k^{(k^{n+1}-2n-k)}(k^n-1)^{2(k^n-1)+1},$$

$$z=k^{k^n(k^{n+1}-2n-k)+n+1}(k^n-1)^{2(k^n-1)+1},$$

其中 $k=2$ 时，$n>1$；$k\geqslant 3$ 时，$n>0$。

类似于定理 1.3.21，有

定理 1.3.23 方程(2)有整数解，至少存在一个 i，$1\leqslant i\leqslant k$，使 x_i 的每一个素因子皆整除 $x_1x_2\cdots x_{i-1}x_{i+1}\cdots x_k$。

到目前为止，人们找到方程(2)的解，除了 $x_1=x_2=\cdots=x_k=z=1$ 外，均是偶数解，于是可提出如下猜想：

没有正奇数 $x_1(>1)$，$x_2(>1)$，…，$x_k(>1)$，$z(>1)$，适合于(2)。

在 6 的(5)式中，设 $x+h+1=y$，(5)中等号左边的 n 变为 1，且设 y 与 x，h 不一定有关系，等号左边由加变成乘，则有

$$x(x+1)\cdots(x+h)=y^n,n>1。\tag{3}$$

由定理 1.3.17 可推得如下结果

定理 1.3.24 不定方程(3)在 $h=2$，和 $h=3$ 时均无非零的整数解。

---

① 柯召，孙琦，关于方程 $\prod_{i=1}^{k}z_i^{n_i}=z^n$，四川大学学报（自然科学版），2(1964)，5—9。

由此，我们可推测有如下猜想：

当 $h>0$ 时，不定方程(3)无非零整数解。

1975 年，艾道斯和舍尔夫里治（Selfridge）分别证明了这一猜想。

上面我们从直角三角形三边的关系式出发，通过一般化，特殊化，类比等各种逻辑组合提出一些问题和猜想，通过问题的探讨与解决，就形成了新的命题和理论，再以新命题为出发点，继续利用上述方法，又提出新的问题和猜想来，这样继续进行下去，就可以演变出许许多多的问题和猜想，从而丰富数学科学的内容。

（四）一条著名的旁注

当我们求得不定方程

$$x^2+y^2=z^2$$

的所有正整数解之后，很自然地使我们想到，当不定方程中元的次数增加 1，又会怎样呢？即求不定方程

$$x^3+y^3=z^3 \tag{1}$$

的正整数解。

我们可以通过试验的方法，看看这个方程是否有正整数解。前 10 个正整数的立方是

1，8，27，64，125，216，243，512，729，1000。

通过观察，不难看出没有一个数可以表示为另外两个立方之和。例如，若 512 可以表示为另外两个立方之和，那么这两个数一定比 512 小，因此在立方表中 512 的前面，最有可能的数就是 243 和 216，但 $243+216=559>512$。而 $216+216=432<512$，所以 512 不是任何两个立方之和。前 20 个正整数的立方是否能找出一个数可以表示为另外两个立方之和？前 20 个正整数的立方除了前面列出的 10 个之外，还有

1331, 1728, 2197, 2744, 3375, 4096, 4913, 5832, 6859, 8000。通过观察，不难看出，也找不到一个数可以表为另外两个立方之和。

用上述方法我们可以验证其他的一个立方数不能表示为另外两个立方之和。当然上述的验证方法还可以再简单一些，比如两个相同数相加，改为 $4096 \div 2 = 2046$，只要查一下立方表没有此数即可。但无论如何，你得化费一定时间去验证。如果你有时间，你可以继续找下去，但总是找不到一个立方数可以表示为另外两个立方之和。

既然对于(1)找不到正整数解，对于不定方程

$$x^4 + y^4 = z^4 \tag{2}$$

又会怎样呢？同样，我们用试验方法，也找不到一个四次方数可以表示为另外两个四次方数之和，即对于方程(2)也没有正整数解。继而考虑不定方程

$$x^5 + y^5 = z^5 \tag{3}$$

的正整数解。我们还是用试验的方法，也找不到一个五次方数可以表示为另外两个五次方之和，即不定方程（3）也没有正整数解。于是，由特殊到一般，我们猜想：

当 $n \geqslant 3$ 时，不定方程

$$x^n + y^n = z^n \tag{4}$$

不存在正整数解。

在历史上，这个猜想称为费尔马大定理。之所以这样称谓，其原因如下。

费尔马于 1601 年出生于法国图卢兹附近，于 1665 年去世于卡斯特尔。他是一个皮革商的儿子，童年是在家庭里受的教育。三十岁时，他在图卢兹当了律师。在当律师期间，他把自己大量的业余时间用在数学研究上。他和同时代的第一流数学家有科学通信关系，使他能够广泛地参与当时的学术活动。他一辈子虽然发表的数学著作不多，但已显示出他是一位有多方面数学成就的

杰出数学家。对于解析几何、微积分以及概率论的创立，都做出了他自己的贡献。特别是，在现代数论的奠基工作中，显示出他有非凡的数学洞察能力。最初吸引费尔马注意的，可能是梅齐利亚克1621年翻译的丢番图《算术》的拉丁文译本。他在钻研这本书时，在上面做了大量旁注。原书有这样一个问题：求一个平方，它是另外两个平方之和。费尔马在这个问题的旁边写道："另一方面，不可能把一个立方表为两个立方之和，把一个四次方表为两个四次方之和，或者，一般来说，一个次数大于2的方幂不可能是两个同次方幂之和。我已经发现了一个确实非常奇妙的证明，但是书的页边太窄了，写不下。"这个命题，后来数学史学家称为"费尔马大定理"，也有的称作"费尔马最后定理"，以示这个定理证明之艰难。实际上，这个命题称为费尔马猜想更确切些。费尔马是否真的给出这个定理的一个完善证明，也许将永远是个谜。很可能，他在写旁注时，有一个证明的想法，而后来，他认识到这个想法是错误的。由于可以肯定断言，他并没有打算把这些旁注公诸于世，因此，他也就不再去修改和去掉这条旁注了。1670年，也就是在他死后五年，这些笔记由他的儿子萨穆埃尔（Samuel）编入《算术》新版发表，从此费尔马大定理才为世人所知，许多人都为了寻求其证明而付出了巨大的努力。然而，人们经过三百多年的努力，只证明了许多特殊情况下的费尔马大定理成立。同时，在证明这个定理的过程中建立了理想数论，对其他数学问题的解决起到了重要的工具作用。直到1993年，英国数学家安德鲁·维尔斯终于在前人研究成果的基础上证明了费尔马大定理，并于1995年将其严格的证明通过权威的美国《数学年刊》公诸于世。

二、长路漫漫

——费尔马大定理的探讨

（一）$n=4$ 的费尔马大定理

证费尔马大定理先从什么地方开始呢？我们可首先考虑它的特殊情况，比如考虑当 $n=3$，4，5 时定理是怎样证明的。这也是解决一般问题常采用的方法，特别是解决难题时，更是不可缺少的思想方法。解决费尔马问题的历史也正是依循这样的一个思考路线。我们先从 $n=4$ 开始。为什么不先从 $n=3$ 开始呢？这是因为解决后者的问题比前者难一些。解决 $n=4$ 的情况，经过简单的变换，还可以利用已讨论过的不定方程 $x^2+y^2=z^2$ 的有关结果。

定理 2.1.1 不定方程

$$x^4+y^4=z^4 \tag{1}$$

无正整数解。

证 设 $u=x$，$v=y$，$w=z^2$，则(1)变成 $u^4+v^4=w^2$。因此求证（1）无正整数解，就转化为求上述不定方程无正整数解。为此，我们只要给出下述定理的证明即可。

定理 2.1.2 不定方程

$$x^4+y^4=z^2 \tag{2}$$

无正整数解。

我们将采用费尔马的无限递降法予以证明。记(2)的所有具有正整数解的集合为 M，然后证明这个集合为空集。如若不然，将导出矛盾。在 M 中，必然有一组使 z^2 的值最小，设这个值为 c^2。使 z 取此值的解可能有 n 组，我们就任取一组，记为 a，b，c。我们还可再构造出一组数 a'，b'，c'，使它们满足 $x^4+y^4=z^2$，但 $c'^2<c^2$，由于 c^2 最小，因而知 M 非空的假设是不对的。故导出(2)没有正整数解。这就是证明该定理的大体思路，具体证明如下。

证 我们假定(2)有一组正整数解 a，b，c，而 c^2 使 z^2 的值最小，且还可认为 a，b 互素。如若不然，就存在一个素数 p，使得 $p|a$，$p|b$，从而 $p^2|c$。于是

$$\left(\frac{a}{p}\right)^4+\left(\frac{b}{p}\right)^4=\left(\frac{c}{p^2}\right)^2,$$

它为(2)提供了另一组解，且相应的值比 c^2 还小，但根据我们的假设，这是不可能的。

a 与 b 不同为奇数，也不同为偶数，它们之中，一为奇数，另一个为偶数，不妨设 a 为偶数。现在我们对(2)有了一组正整数解：

$$(a^2)^2+(b^2)^2=c^2,$$

其中 $(a^2,\ b^2)=1$，a^2 为偶数，b^2 为奇数，因此存在整数 m，n，两数互素，且不同为奇数，使

$$\begin{cases} a^2=2mn, \\ b^2=m^2-n^2, \\ c=m^2+n^2。\end{cases} \tag{3}$$

(3)中的 m，n 为一奇一偶，令 $n=2q$，那么由(3)得

$$a^2=4mq \text{ 或}$$

$$\left(\frac{a}{2}\right)^2=mq。\tag{4}$$

我们可证明 m，q 互素，进而推知 m 和 q 均为平方数，设 $m=c'^2$，$q=a'^2$，不难验证 c'，a'互素，且 c'为奇数。将式

$$n=2q=2a'^2, m^2-n^2=b^2, m=c'^2$$

代入下式

$$n^2+(m^2-n^2)=m^2,$$

就有

$$(2a'^2)^2+b^2=(c'^2)^2。$$

于是我们求出另一组勾股数。

由于 $(2a'^2,b)=1$，故知 $2a'^2$，b，c'^2 是方程 $x^2+y^2=z^2$ 的一组整数解，其中 $2a'^2$ 是偶数。于是根据我们已经证明的结论，存在整数 M 和 N，其中 $(M,N)=1, M\not\equiv N(\bmod 2)$，使

$$\begin{cases}2a'^2=2MN,\\ b=M^2-N^2,\\ c_1^2=M^2+N^2。\end{cases}\tag{5}$$

因此，有 $a'^2=MN$，$(M,N)=1$。两个互素的整数，当且仅当它们都是平方数时，它们的乘积也是平方数。所以存在 a_1 和 b_1，使

$$M=a_1^2,\quad N=b_1^2。$$

由(5)的第三式，得

$$c_1^2=(a_1^2)^2+(b_1^2)^2$$

或

$$a_1^4+b_1^4=c_1^2。$$

这就导出了(2)的另一组解，且有

$$c_1^2=m\leqslant m^2<m^2+n^2=c\leqslant c^2。$$

这是不可能的，因为 c^2 是最小的。从而定理 2.1.2 得证。（证完）

推论 不定方程

$$x^n+y^n=z^n,\quad n=4k,\quad k\text{ 为正整数},\tag{6}$$

无正整数解。

证 将不定方程

$$x^n + y^n = z^n$$

改写为

$$(x^k)^4 + (y^k)^4 = (z^k)^4,$$

令 $x^k = u$，$y^k = v$，$z^k = w$，则得

$$u^4 + v^4 = w^4,$$

而此方程由定理 2.1.1 知无正整数解，故(6)无正整数解。

（二）关于 $n=3$ 的欧拉证明

定理 2.2.1 不定方程

$$x^3 + y^3 = z^3 \tag{1}$$

无正整数解。

为了证明(1)，先介绍欧拉的证法，在下一节将给出另一个初等证法。不难证明，如果 n 等于某一个整数 r 时，费尔马大定理成立，那么 n 等于 r 的倍数时定理亦成立。因为任何整数或者被 4 或者被奇素数除尽，所以要证明费尔马大定理成立，只要证明 n 等于奇素数的情况即可。欧拉应用了费尔马的无穷递降法，对于第一个奇素数的情况，证明了费尔马大定理，尽管证明中还存在一定的缺陷，但欧拉的思想还是很有启发性的。

欧拉是他那个时代最伟大的数学家之一，他一生中留下了浩瀚的数学著作，有八百多篇，史学家称他为“数学英雄”。欧拉于 1707 年生于瑞士的巴塞尔附近，于 1783 年去世。他的父亲是个数学爱好者，是欧拉的第一个数学老师。当欧拉还是中学生时，他就利用业余时间到大学去旁听数学。1766 年他的双眼完全失明了，但欧拉从没有中断科学研究工作，还孜孜不倦，顽强学习，刻苦钻研，在这之后竟发表了四百多篇论文。无论是应用数学，还是纯粹数学，在数学各个分支几乎都有他的贡献。在他的著作中不仅给人们创造了数学的新成果，而且还留下了数学家

发现新定理的思想方法。他总是下功夫把有关的归纳证据细心地、详尽地、有条理地写出来。他的大胆猜测和巧妙证明，常常成为启发人们灵感的重要源泉。欧拉对于费尔马大定理的贡献，也同样凝聚了他的伟大发现的思想。

至于欧拉是否证明了费尔马大定理 $n=3$ 的情况，在费尔马大定理研讨的历史上是存在着争议的。普遍认为，欧拉证明了费尔马大定理 $n=3$ 的情况，但证明并不完善。事实上，在他给出的证明中，包含着一个严重错误，而他当时并没有发现。用最直接的方法改正欧拉的证明是麻烦的。然而，正如我们将要讲的，欧拉的证明可以用不太直接的方法来改进，即利用欧拉用来证明费尔马其他命题所用到的一些方法来讨论。

1. 欧拉关于 $n=3$ 的证明

欧拉证明(1)式的基本方法是费尔马的无穷递降法。他证明了如果可以找到正整数使（1）成立，则可以找到更小的正整数满足(1)式。于是可以找这样三元正整数组的一个连续递降的序列来，而这是不可能的，所以没有这样的正整数满足(1)式。下面就来具体叙述证明过程。

证 设有正整数 x，y，z 满足(1)式，不妨假设 x，y，z 两两互素，即 $(x,y)=(x,z)=(y,z)=1$，如若不然，任何整除 x，y，z 中两个的因子必整除第三个，因此可以约去其公因子。特别地，x，y，z 中至多有一个偶数，从而其中恰有一个偶数。

先假设 x，y 是奇数，z 是偶数，则 $x+y$，$x-y$ 均为偶数，分别设为 $2p$，$2q$，从而

$$x=\frac{1}{2}(2p+2q)=p+q,$$

$$y=\frac{1}{2}(2p-2q)=p-q。$$

当 $x^3+y^3=(x+y)(x^2-xy+y^2)$ 用 p，q 来表达时，有

$$2p[(p+q)^2-(p+q)(p-q)+(p-q)^2]$$
$$=2p(p^2+3q^2),$$

其中 p，q 是一奇一偶(因为 $p+q$，$p-q$ 均为奇数)，且它们是互素的(因为任何公因子必为 $x=p+q$ 与 $y=p-q$ 的公因子，从而为1)。进而可以假设 p，q 大于零（因为如果 $x<y$，则交换 x 与 y，从而得 $q>0$；如果 $x=y$，这是不可能的，因为此时必有 $x=y=1$，从而 $z^3=2$)。因此，由

$$x^3+y^3=z^3,\ x,\ y \text{ 均为奇数} \tag{2}$$

可推出，存在互素的一奇一偶的正整数 p，q，使得

$$2p(p^2+3q^2)=\text{立方数。}$$

如果 z 是奇数，而 x 或 y 为偶数，不妨假设 x 为偶数，将有同样的结论。在这种情况下，将(2)中的 y^3 移到等号右边，得

$$x^3=z^3-y^3=(z-y)(z^2+zy+y^2)\text{。}$$

设 $z-y=2p$，$z+y=2q$，从而 $z=q+p$，$y=q-p$，代入上式，得

$$x^3=2p[(q+p)^2+(q+p)(q-p)+(q-p)^2]\text{。}$$

从而，得到相同的结论

$$2p(p^2+3q^2)=\text{立方数，}$$

其中 p，q 为互素的一奇一偶的正整数。

下面证明 $2p$ 和 p^2+3q^2 是互素的，且它们的积 $2p(p^2+3q^2)$ 是立方数，只有它们每一个是立方数。因为 p，q 为一奇一偶，所以 p^2+3q^2 是奇数，且 $2p$，p^2+3q^2 的任何一个公因子必为 p，p^2+3q^2 的一个公因子，从而必是 p，$3q^2$ 的公因子。因为 p，q 互素，故其仅可能的公因子为3。但如果3整除 p，则显然3整除 p^2+3q^2，从而 $2p$，p^2+3q^2 不互素。所以现在分成两种情况证明：i）3不整除 p，从而 $2p$，p^2+3q^2 互素；ii）3整除 p。先证 i)，ii）由 i）作适当修改即得。

假设3不整除 p,从而 $2p$, p^2+3q^2 均为立方数。利用公式

$$(a^2+3b^2)(c^2+3d^2)=(ac-3bd)^2+3(ad+bc)^2$$

可求出形如 p^2+3q^2 的立方：

$$
\begin{aligned}
(a^2+3b^2)^3 &= (a^2+3b^2)[(a^2-3b^2)^2+3(2ab)^2] \\
&= [a(a^2-3b^2)-3b(2ab)]^2 \\
&\quad +3[a(2ab)+b(a^2-3b^2)]^2 \\
&= (a^3-9ab^2)^2+3(3a^2b-3b^3)^2。
\end{aligned}
$$

这就是说，求形如 p^2+3q^2 的立方数的一种方法就是找数 a，b，令

$$p=a^3-9ab^2,\quad q=3a^2b-3b^3,$$

使得 $\quad p^2+3q^2=(a^2+3b^2)^3$。

欧拉的证明中需要补充的证明就是：若 p^2+3q^2 是一个立方数，则必存在 a，b，使得 p，q 由上述等式给出。关于这一结论我们将在 4. 中给出证明。欧拉利用了一个错误的方法证明了这一结论，这将在 2. 中叙述。现在我们暂且承认这一结论，继续往下证。

p 与 q 的表达式可以分解为

$$p=a(a-3b)(a+3b),q=3b(a-b)(a+b)。$$

当然 a 与 b 是互素的，这是因为 a 与 b 的任何一个公因子整除 p 与 q。进而

$$2p=2a(a-3b)(a+3b)\ \overline{\overline{\gamma}}\ 立方数,$$

a 与 b 之奇偶必不同，否则 p 与 q 必同为偶数。于是 $a-3b$，$a+3b$ 均为奇数，且 $2a$，$a\pm3b$ 仅有可能的公因子为 a，$a\pm3b$ 的公因子，即 a，$\pm3b$ 的公因子。类似地，$a+3b$ 与 $a-3b$ 之公因子必为 a 与 $3b$ 之公因子。这就是说，仅有可能的公因子为 3。但 3 不整除 a，否则 3 将整除 p，此与假设矛盾。因此 $2a$，$a-3b$，$a+3b$ 是互素的，且每一个必为立方数。设 $2a=\alpha^3$，$a-3b=\beta^3$，$a+3b=\gamma^3$，则 $\beta^3+\gamma^3=2a=\alpha^3$，从而给出了(1)的一个解，且比原来的小。

更确切地说，$\alpha^3\beta^3\gamma^3=2a(a-3b)(a+3b)=2p$，当 z 是偶数时，这是 z^3 的一个因子。而当 x 是偶数时，它是 x^3 的因

子。无论怎样，$\alpha^3\beta^3\gamma^3$ 小于 z^3。由于$(-\alpha)^3=-\alpha^3$，且负数的立方移至方程的另一边后就变成了正数的立方，且总是得到形如 $x_0^3+y_0^3=z_0^3$ 的方程，其中 x_0，y_0，z_0 全是正数，且 $z_0^3<z^3$，所以不用避免 α，β，γ 为负。因此在 3 不整除 p 时，我们已经完成了解的递降。

现在考虑 ii)，即 $3\mid p$。设 $p=3s$，则 $3\nmid q$，于是 $2p(p^2+3q^2)=3^2\cdot 2s(3s^2+q^2)$，易见 $3^2\cdot 2s$ 与 $3s^2+q^2$ 是互素的，从而它们均为立方数。易见其因子是互素的，从而 $2b=\alpha^3$，$a-b=\beta^3$，$a+b=\gamma^3$。进而 $\alpha^3=2b=r^3-\beta^3$。同上面完全一样，由此得出形如 $x_0^3+y_0^3=z_0^3$ 的方程，其中 $z_0^3<z^3$。

综合 i)，ii) 可知，由一个立方数是另外两个立方数之和就可推出存在更小的一个立方数是另外两个立方数之和。这是不可能的。从而定理 2.2.1 得证。(证完)

2. 根式环

欧拉为了建立形如 p^2+3q^2 的立方数所需要的命题就研究了形如 $a+b\sqrt{-3}$（a，b 为整数）的数组成的数系 R，R 与整数系很相似。设 $a_1+b_1\sqrt{-3}\in R$，$a_2+b_2\sqrt{-3}\in R$，则

i) $(a_1+b_1\sqrt{-3})+(a_2+b_2\sqrt{-3})=(a_1+a_2)+$
$+(b_1+b_2)\sqrt{-3}\in R$,

ii) $(a_1+b_1\sqrt{-3})-(a_2+b_2\sqrt{-3})=(a_1-a_2)+$
$+(b_1-b_2)\sqrt{-3}\in R$,

iii) $(a_1+b_1\sqrt{-3})(a_2+b_2\sqrt{-3})=a_1a_2+a_1b_2\sqrt{-3}+$
$+b_1a_2\sqrt{-3}+a_1b_2(-3)=(a_1a_2-3a_1b_2)+(a_1b_2+$
$+b_1a_2)\sqrt{-3}\in R$,

iv) $1\cdot(a_1+b_1\sqrt{-3})=a_1+b_1\sqrt{-3}$,

而且 R 中的数 $a+b\sqrt{-3}$ 亦适合结合律、交换律及分配律。所

以 R 为一个具有单位元的交换环。

计算数 $a+b\sqrt{-3}$的目的就在于简化 p^2+3q^2 成立方数的充分条件。不利用

$$(a_1^2+3b_1^2)(a_2^2+3b_2^2)$$
$$=(a_1a_2-3a_1b_2)^2+3(a_1b_2+b_1a_2)^2,$$

我们可如下讨论。分解

$$p^2+3q^2=(p+q\sqrt{-3})(p-q\sqrt{-3})。$$

易见，如果其中一个因子是一个立方数，比如说

$$p+q\sqrt{-3}=(a+b\sqrt{-3})^3,$$

则 $p-q\sqrt{-3}$必为 $a+b\sqrt{-3}$之共轭即 $a-b\sqrt{-3}$之立方，即

$$p-q\sqrt{-3}=(a-b\sqrt{-3})^3。$$

因此，由乘法的交换律知

$$(p+q\sqrt{-3})(p-q\sqrt{-3})$$
$$=[(a+b\sqrt{-3})(a-b\sqrt{-3})]^3,$$

即 $p^2+3q^2=(a^2+3b^2)^3$ 。换言之，要求形如 p^2+3q^2 的立方数只需找 a，b，使 $p+q\sqrt{-3}=(a+b\sqrt{-3})^3$ ，即

$$p+q\sqrt{-3}=a^3+3a^2b\sqrt{-3}+3ab^2(-3)$$
$$+b^3(-3)\sqrt{-3}。$$

因此，p^2+3q^2 是一个立方数当且仅当存在整数 a，b，使得 $p=a^3-9ab^2$，$q=3a^2b-3b^2$。这正是上一节中所要证明的充分条件。

欧拉在他著的“代数”这一部分中，严重地混淆了必要充分条件，并且很难确定他要想说什么。在他所举的例子中，似乎大部分是想处理充分条件，从 a，b 开始找 p，q，但他的确有些错误的结论。例如他说：“当 x^2+cy^2 是一个立方数时，其两个既约因子，即 $x+y\sqrt{-c}$，$x-y\sqrt{-c}$必为立方数，这是因为它们是互素的（如果 x，y 无公共因子的话）”，尽管他没有给出 $x+y\sqrt{-c}$及$x-y\sqrt{-c}$均为立方数的证明。在这本书中同一段的一个命题中，他非常含混地说：“若 ax^2+cy^2 不能分解成两

个有理因子，则不存在异于这里给出的解。”即除了存在整数 p，q，使得 $x\sqrt{a}+y\sqrt{-c}=(p\sqrt{a}+q\sqrt{-c})^3$ 之外，没有别的方法，使 ax^2+cy^2 成为立方数了。关于这一事实的证明仅有一点暗示，就是由上边给出的类似讨论所得出的结论，即，若 AB 是立方数，A，B 互素，则 A，B 均为立方数。如果 A，B 均为整数，则这是一个可以证明的定理。但这个证明对于 A，B 为形如 $p+q\sqrt{-3}$的数其结论就不一定成立了。事实上，欧拉的结论对于 $a=1$，$c=3$ 是对的，尽管其证明与整数的情形不太一样。但是对于有一些 a，c，这个命题就不成立了。

与 x^2+cy^2 等于立方数的讨论紧密相关的是欧拉给出了 x^2+cy^2 等于平方数的一个相当完全的讨论。他说：“如果两个数之积，比如 pq 是一个平方数，则必有 $p=r^2$，$q=s^2$，或者 $p=mr^2$，$q=ms^2$。这就是说，每个因子刚好是一个平方数乘上相同一个数。”如果这里的“数”是指整数，这是成立的。但欧拉立即就将此结论用到了形如 $x+y\sqrt{-c}$的数上面了。欧拉似乎证明了：若 x^2+cy^2 是一个平方数，且 x，y 是互素的整数，则存在整数 a，b，使 $x+y\sqrt{-c}=(a+b\sqrt{-c})^2$。从这一点看，欧拉关于平方数的讨论比关于立方数的讨论完全些。这个“证明”用了标准的丢番图方法：将 x^2+cy^2 的平方根写成 $x+(p/q)y$ 的形式，其中 p 与 q 是这样定义的整数，然后简化：

$$\left(x+\frac{p}{q}y\right)^2=x^2+cy,$$

$$\frac{2p}{q}xy+\frac{p^2}{q^2}y^2=cy^2,$$

$$\frac{2p}{q}\frac{x}{y}=\frac{cq^2-p^2}{q^2},$$

$$\frac{x}{y}=\frac{cq^2-p^2}{2pq}。$$

“但正如 p，q 是互素的一样，x，y 是互素的，从而 $x=cq^2-p^2$，$y=2pq$。”所以 $-x+y\sqrt{-c}=(p+q\sqrt{-c})^2$。欧拉说这

个推导“加强了上述方法的正确性。”但是 p 与 q 互素这一自然的假设并不能推出 cq^2-p^2 与 $2pq$ 是互素的，从而也就不一定得出欧拉的结论：

$$x=cq^2-p^2,\quad y=2pq。$$

事实上，例如 $49=2^2+5\cdot3^2$，说明上述讨论是不对的，因为方程

$$\begin{cases}2=5q^2-p^2,\\3=2pq\end{cases}$$

就根本没有有理解（这两条相交双曲线相交于

$$p=\sqrt{\frac{5}{2}},\quad q=\sqrt{\frac{9}{10}}\text{以及 } p=-\sqrt{\frac{5}{2}},\quad q=-\sqrt{\frac{9}{10}}\text{）。}$$

欧拉已经注意到其方法有一定的局限性。他为了求解 $2x^2-5=$ 立方数，他的方法导致

$$x\sqrt{2}+\sqrt{5}=(a\sqrt{2}+b\sqrt{5})^3=a^3 2\sqrt{2}$$
$$+3a^2b2\sqrt{5}+3ab^2 5\sqrt{2}+b^3 5\sqrt{5},$$

从而 $x=2a^3+15ab^2$，$1=6a^2b+5b^3$。而后一方程得 $b=\pm1$，且 $6a^2+5b^2=\pm1$，而这是不可能的。于是欧拉的方法就指出 $2x^2-5$ 不可能是一个立方数。但是，当 $x=4$ 时，有

$$2x^2-5=2\cdot16-5=27=3^3。$$

就是一个立方数。为什么对有的例子行，有的例子不行呢？欧拉最初的答复只是强调研究其数学基础是非常重要的。很自然可以推测，欧拉承认了他的方法在这一点上不完备，并且建议研究解决这一问题。从他的书后面的两节中清楚说明了欧拉认识到了问题的症结就在于 $2x^2-5y^2$ 中的减号以及 Pell 方程 $x^2-10y^2=1$ 除平凡解 $x=\pm1$，$y=0$ 外，仍有其他解，这一相关的事实。无需细究，可以说欧拉已经认为他的方法，在加号的情况下是适用的。但我们可举出反例，对于例子 $49=2^2+5\cdot3^2$，其方法是不适用的，尽管是加号，且其 Pell 方程 $x^2+5y=1$ 只有平凡解 $x=\pm1$，$y=0$。

欧拉在 1753 年 8 月 4 日给哥德巴赫的信中声称他能证明费

尔马大定理 $n=3$ 的情形，但是他发表的证明仅仅在其《代数》中的一处。关于这个有错误的证明的一个合理的推测是：他的原始证明用了一个想象的论点去证明，若 $x^2+3y^2=$ 立方数，则 $x=a^3-9ab^2$，$y=3a^2b-3b^3$，稍后，他有了一个非常优美但错误的想法去证明这点，即他用了：若 $(x+y\sqrt{-3})(x-y\sqrt{-3})$ 是一个立方数，且若因子 $x+y\sqrt{-3}$ 和 $x-y\sqrt{-3}$ 是互素的，则这两个因子本身也必定是立方数。不论此推测正确与否，利用欧拉在他早期工作中所用的思想可以证明关于形如 $x^2+3y^2=$ 立方数这一点是确信无疑的。

3. 关于两平方数之和

欧拉在 1747 年给哥德巴赫的信中宣布他证明了费尔马的一个定理：每一个形如 $4n+1$ 的素数是两个平方数之和。欧拉宣称他的主要目的是证明费尔马的另一个定理，即每个数能够写成四个或者少于四个平方数之和。但这后一个定理难住了欧拉，直到 1770 年，拉格朗日（Lagrange, Joseph Louis, 法国人，1736—1813）证明了这一定理，欧拉才给出一个简化证明。但是欧拉所证明的两个平方数之和的定理是非常重要的，特别是，欧拉证明这个定理所用的技巧也能够使他证明关于形如 x^2+3y^2 的数的一些基本事实，并且如下节所述，这个技巧也能用于证明形如 x^2+3y^2 的立方数的一些证明费尔马大定理在 $n=3$ 时所用的事实。

欧拉关于每个形如 $4n+1$ 的素数是两个平方数之和的证明并不很长，而且也非常初等。

i. 两个由两个平方数之和组成的数之积还是两个平方数之和。

此结论可由公式

$$(a^2+b^2)(c^2+d^2)=(ac-bd)^2+(ad+bc)^2$$

推出。

ii. 如果一个由两个平方数之和组成的数可以被一个由两个平方数之和组成的素数除尽，则商为两个平方数之和。

例如，假定 a^2+b^2 可被素数 p^2+q^2 除尽，则 p^2+q^2 能除尽

$$(pb-aq)(pb+aq)=p^2b^2-a^2q^2=p^2b^2+p^2a^2$$
$$-p^2a^2-a^2q^2=p^2(a^2+b^2)-a^2(p^2+q^2)。$$

因为 p^2+q^2 为素数，所以它整除 $pb-aq$ 或 $pb+aq$。假定 p^2+q^2 整除 $pb+aq$，则由 $(a^2+b^2)(p^2+q^2)=(ap-bq)^2+(aq+bp)^2$ 得 p^2+q^2 必定也整除$(ap-bq)^2$ 。所以此方程可被 p^2+q^2 整除。所以$\dfrac{a^2+b^2}{p^2+q^2}$是两个平方数之和。若 p^2+q^2 整除 $pb-aq$，则由 $(a^2+b^2)\cdot(q^2+p^2)=(aq-bp)^2+(ap+bq)^2$ 可作如上同样的讨论。

iii. 若一个由两个平方数之和组成的数可被一个不是由两个平方数之和组成的数整除，则其商有一因子不是由两个平方数之和组成。

这恰好是第 2 种情况的反面。假定 x 除以 a^2+b^2，其商的素因子分解为 $p_1p_2\cdots p_n$，则

$$a^2+b^2=xp_1p_2\cdots p_n。$$

若所有的因子 p_1，p_2，…，p_n 都能表示成两个平方数之和，那么 a^2+b^2 相继能由 p_1，p_2，…，p_n 除尽。由 ii. 知，每一个商，包括 x 在内，均是两个平方数之和。所以若 x 不是两个平方数之和，则 p_1，p_2，…，p_n 中必定有一个不是两个平方数之和。

iv. 若 a 和 b 是互素的，则 a^2+b^2 的每个因子是两个平方数之和。

假定 x 是a^2+b^2 的一个因子，$a=mx\pm c$，$b=nx\pm d$，$|c|\leqslant\dfrac{1}{2}x$，$|d|\leqslant\dfrac{1}{2}x$。因为

$$a^2+b^2=m^2x^2\pm 2mxc+c^2+n^2x^2\pm 2nxd+d^2$$
$$=Ax+(c^2+d^2)$$

可被 x 除尽，c^2+d^2 必定能被 x 整除。假定 $c^2+d^2=yx$。如果 c 和 d 有任何大于 1 的公因子，则这个因子不能整除 x，因为否则它将整除 a 和 b，与假设矛盾。所以方程 $c^2+d^2=yx$ 能被 c 和 d 的最大公因子的平方整除，得到一个方程 $e^2+f^2=zx$，因为

$$zx=e^2+f^2\leqslant c^2+d^2\leqslant\left(\frac{1}{2}x\right)^2+\left(\frac{1}{2}x\right)^2=\frac{1}{2}x^2,$$

故有 $z\leqslant\frac{1}{2}x$。若 x 不是两个平方数之和，则由（3）知，存在 z 的一个因子 w，不能被写成两个平方数之和。这将导致一个无穷递降，由一个不是两个平方数之和，而是两个互素平方数之和的因子数 x，降到与它具有相同性质的更小的数 w。所以 x 必定是两个平方数之和。

v. 每个形为 $4n+1$ 的素数都是两个平方数之和。

欧拉首先给出这个结论一个含糊的证明，两年以后，也就是在 1749 年又给出这个结论如下的一个优美的证明。

若 $p=4n+1$ 是素数，则由费尔马定理知，1，2^{4n}，3^{4n}，4^{4n}，…，$(4n)^{4n}$ 中的每个数都具有 $mp+1$ 的形式〔后一项 p^{4n} 可被 p 整除，$(p+1)^{4n}$，…，$(2p-1)^{4n}$ 也都具有 $mp+1$ 的形式，m 为任意数，依次类推〕，所以差 $2^{4n}-1$，$3^{4n}-2^{4n}$，…，$(4n)^{4n}-(4n-1)^{4n}$ 都可被 p 整除。这些差中的每一个都可分解为两个因子的乘积：

$$a^{4n}-b^{4n}=(a^{2n}+b^{2n})(a^{2n}-b^{2n})。$$

由 p 为素数知，p 必定整除上面等式右边的两个因子中的一个。若对所有的这 $4n-1$ 个差数，p 都能整除第一个因子，则由任何一个数和它的后继数互素这一事实以及断言（4）知，p 是两个平方数之和，由此可知 p 不能整除这所有的 $4n-1$ 个数：

$$2^{2n}-1,\ 3^{2n}-2^{2n},\ \cdots,\ (4n)^{2n}-(4n-1)^{2n}\ 。$$

因为，若 p 整除所有这 $4n-1$ 个数，则 p 将整除由这 $4n-1$ 个数所形成的 $4n-2$ 个差数，同理也整除这些差的 $4n-3$ 个差数，依次类推。由基础代数可证明序列 1^k，2^k，3^k，4^k，…的 k 阶差是常数且等于 $k!$ （见表所示）。所以序列 1，2^{2n}，3^{2n}，4^{2n}，…的 $2n$ 阶差都等于 $(2n)!$，不能被 $p=4n+1$ 整除。如果 p 整除开始的 $4n-1$ 个一阶差 $2^{2n}-1$，$3^{2n}-2^{2n}$，…，$(4n)^{2n}-(4n-1)^{2n}$，则 p 也整除开始的 $4n-2n$ 个 $2n$ 阶差，实际不然，所以 p 不能整除这所有的 $4n-1$ 个一阶差。

表　x^k 的差

$k=1$	1	2	3	4	5	6	7	…	
1 阶差	1	1	1	1	1	1	…		
2 阶差	0	0	0	0	0	…			
$k=2$	1	4	9	16	25	36	49	…	
1 阶差	3	5	7	9	11	13	…		
2 阶差	2	2	2	2	2	…			
3 阶差	0	0	0	0	…				
$k=3$	1	8	27	64	125	216	343	512	…
1 阶差	7	19	37	61	91	127	169	…	
2 阶差	12	18	24	30	36	42	…		
3 阶差	6	6	6	6	6	…			
4 阶差	0	0	0	0	…				
$k=4$	1	16	81	256	625	1296	2401	4096	…
1 阶差	15	65	175	369	671	1105	1695	…	
2 阶差	50	110	194	302	434	590	…		
3 阶差	60	84	108	132	156	…			
4 阶差	24	24	24	24	…				
5 阶差	0	0	0	…					

1658 年，费尔马在一封信中谈到他有无可辩驳的事实说明

下面的结论是对的：(i)形为 $4n+1$ 的素数都是两个平方数之和；(ii) 形为 $3n+1$ 的每个素数具有形式 a^2+3b^2；(iii) 每个形为 $8n+1$ 或 $8n+3$ 的素数具有形式 a^2+2b^2；(iv) 每一个数是 3 个或少于 3 个的三角形数之和①，4 个或少于 4 个平方数之和，5 个或少于 5 个五角形数② 之和等等。欧拉对于断言(iv)非常有兴趣，特别是每个数是 4 个或少于 4 个平方数之和这一断言。非常自然地，欧拉试图用他证明第(i)个断言所用的技巧去证明其余的断言。他发现，(iii)与(iv)他是无能为力的，但是(ii)可以用与(i)几乎相同的方法去证明。

在 (ii) 和 (i) 的证明中，主要区别在于对于素数 2 的处理上。注意上面的结论(4)对于形为 a^2+3b^2 的表示是错误的，因为 $1^2+3\cdot1^2$ 可被 2 除尽，尽管 2 不为形为 a^2+3b^2 的数。同时，还需要注意到，若结论 (4) 被用到形为 a^2+3b^2 的表示中，则不等式

$$zx\leqslant\left(\frac{1}{2}x\right)^2+\left(\frac{1}{2}x\right)^2=\frac{1}{2}x^2$$

变成

$$zx\leqslant\left(\frac{1}{2}x\right)^2+3\left(\frac{1}{2}x\right)^2=x^2,$$

所以 $z\leqslant x$。但在证明递降时需用到严格的不等式 $z<x$，这儿却得不到。但是，若 x 是奇数，则不等式 $|c|\leqslant\frac{1}{2}x$，$|d|\leqslant\frac{1}{2}x$ 是严格的不等式 $|c|<\frac{1}{2}x$，$|d|<\frac{1}{2}x$，则所需要的不等式 $z<x$ 可以得到。对于表示成形为 a^2+3b^2 的数类似于结论(4)的证明可如下给出。

① 两个可以被写成形为 a^2+3b^2 的数的乘积，也具有形式

---

① 数 1，3，6，10，…，$\frac{n(n+1)}{2}$ 叫三角形数。

② 数 1，5，12，22，…，$\frac{n(3n-1)}{2}$ 叫五角形数。

a^2+3b^2。

此断言可由

$$(a^2+3b^2)(c^2+3d^2)=(ac-3bd)^2+3(ad+bc)^2$$

推出。

② 如果一个形为 a^2+3b^2 的数可被 2 整除，则它也可以被 4 整除，被 4 除所得的商具有形式 c^2+3d^2。

若 a，b 一奇一偶，则 a^2+3b^2 不能被 2 整除。若 a，b 都是偶数，则 a^2+3b^2 可被 2^2 整除，且商具有形式 c^2+3d^2，其中 $c=\frac{1}{2}a$，$d=\frac{1}{2}b$。若 a，b 都是奇数，则 $a=4m\pm1$，$b=4n\pm1$（这里的 m，n 和符号都是适当选取），所以 $a+b$ 或 $a-b$ 是可以被 4 整除的。若 $a+b$ 被 4 整除，则

$$\begin{aligned}4(a^2+3b^2)&=(1^2+3\cdot1^2)(a^2+3b^2)\\&=(a-3b)^2+3(a+b)^2\end{aligned}$$

可被 4^2 整除〔因为 $a-3b=(a+b)-4b$〕,所以 $(a^2+3b^2)/4$ 具有形式 c^2+3d^2。若 $a-b$ 被 4 整除，则可以 $4=(-1)^2+3\cdot1^2$ 代替 $4=1^2+3\cdot1^2$,得到与上面相同的结论。

③ 若形为 a^2+3b^2 的数可被形为 p^2+3q^2 的素数整除，则其商可被写成形式 c^2+3d^2。

由

$$\begin{aligned}(pb-aq)(pb+aq)&=p^2b^2+3q^2b^2-3q^2b^2-a^2q^2\\&=b^2(p^2+3q^2)-q^2(a^2+3b^2)\end{aligned}$$

可被 p^2+3q^2 整除，又 p^2+3q^2 是素数，得知 $pb-aq$ 或 $pb+aq$ 可被 p^2+3q^2 整除。所以

$$\begin{aligned}(p^2+3q^2)(a^2+3b^2)&=[p^2+3(\pm q)^2](a^2+3b^2)\\&=(pa\mp3qb)^2+3(pb\pm aq)^2\end{aligned}$$

能被 $(p^2+3q^2)^2$ 整除（当符号适当选取时）。

故　$(a^2+3b^2)/(p^2+3q^2)$ 有所求之形式。

④ 若形为 a^2+3b^2 的数有一奇因子不具有该形式，则商也

有一不具有此形式的奇因子。

设 $xy=a^2+3b^2$，其中 x 为奇数。若 y 是偶数，则由断言②知，它可被 4 整除，且 $x(y/4)=c^2+3d^2$，这种过程可被重复进行下去，直到 $y/4^K$ 是奇数为止。所以 $y=p_1p_2\cdots p_n$，其中 p_i 是 y 或一个奇素。若 y 的这种分解中的所有奇素数可被写成形为 c^2+3d^2，则 $xy=a^2+3b^2$，能被每一个 p_i 整除。且由断言 2′，3′知，x 可被写成形式 c^2+3d^2，所以 x 不具有此种形式，那么 y 一定有一奇因子不具有此种形式。

⑤ 若 a，b 互素，则 a^2+3b^2 每一奇因子都具有形式 c^2+3d^2。

假定 x 是 a^2+3b^2 的一个奇因子，$a=mx\pm c$，$b=nx\pm d$，其中 $|c|<\frac{1}{2}x$，$|d|<\frac{1}{2}x$，则 c^2+3d^2 可被 x 整除。设 $c^2+3d^2=xy$，其中 $y<x$，则没有大于 1 的 c 和 d 的公因子能整除 x。因为若不然，则与 a 和 b 互素矛盾。所以 $c^2+3d^2=xy$ 能被 c 和 d 的最大公因子除尽，给出 $e^2+3f^2=xz$，其中 e，f 互素。若 x 不具有形式 a^2+3b^2，则由结论④知，z 有一奇因子 w 不具有此种形式。所以若存在一个奇数 x，它是 a^2+3b^2 的一个因子且 x 本身不具有形式 c^2+3d^2，则意味着一个更小的具有和 x 同样性质的 w 存在，由其无穷递减性，可得所求结论。

除 3 以外，每个素数都具有形式 $3n+1$ 或 $3n+2$。具有形式 $3n+2$ 的数一定不具有形式 a^2+3b^2。因为如果 a^2+3b^2 不能被 3 整除，则 a 必定不能被 3 整除，$a=3m\pm1$，且 a^2+3b^2 是 3 的倍数加 1，这样由断言⑤知，一个可以整除形为 a^2+3b^2 的数的奇素数（a 和 b 互素）不具有形式 $3n+2$。若 a 和 b 不互素，则

$$a^2+3b^2=d^2(e^2+3f^2),$$

其中 d 是 a 和 b 的最大公因子，e 和 f 互素。这样一个形为 a^2+3b^2 的数能被写成一个没有形为 $3n+2$ 的奇素因子的数的平方倍

数。对于 a^2+3b^2 的偶素因子，即此数整除 a^2+3b^2，且是 2 的倍数，由断言②知，它是 4 的幂，故为一个平方数。所以一个数具有形式 a^2+3b^2 的一个必要条件是由它所包含的最大平方数所得的商不包含形为 $3n+2$ 的素因子。此条件也是充分的，这只需要证明：

⑥ 形为 $3n+1$ 的每个素数都具有形式 a^2+3b^2。

同前所证，由费尔马定理知，素数 $p=3n+1$ 整除由数 1，2^{3n}，3^{3n}，…，$(p-1)^{3n}$ 所得的 $p-2$ 个差。这些差数每一个可被分解成

$$a^{3n}-b^{3n}=(a^n-b^n)(a^{2n}+a^nb^n+b^{2n}),$$

且第二个因子可被写成（因为 a 或 b 是偶数）：

$$A^2+A(2B)+(2B)^2=(A+B)^2+3B^2,$$

其中 A 和 B 互素。所以由断言⑤知，p 除必定具有形式 c^2+3d^2 外，它整除由数 1，2^n，3^n，…，$(p-1)^n$ 所得的 $p-2$ 个差。同前所证知 p 将整除 $n!$。这是不可能的。所以 p 必具有形式 c^2+3d^2。

若把上述论证应用到形为 a^2+2b^2 的表示上去，可以证明一个数具有形式 a^2+2b^2 的一个必要条件是它被它所含的最大平方数整除所得的商，不包含形为 $8n+5$ 或 $8n+7$ 的素因子。这个条件也是充分的，这只需要证明形为 $8n+1$ 或 $8n+3$ 的所有素数可被写成 a^2+2b^2。正是这最后一步，欧拉没能证明，首先给出证明的是拉格朗日。

4. 一个引理的证明

欧拉计算形为 $a+b\sqrt{-c}$ 的数，是基于公式

$$(x^2+cy^2)(u^2+cv^2)=(xu-cyv)^2+c(xv+yu)^2$$

此公式在上面多次出现。简单地说，若整数 A 是整数 B 和 C 的乘积，且若 B 和 C 能写成形为 a^2+cb^2，即 $B=x^2+cy^2$，$c=u^2+cv^2$，则 A 也可以写成此种形式（利用公式 $a+b\sqrt{-c}=(x+$

$y\sqrt{-c})(u+v\sqrt{-c})$ 来定义 a 与 b)。

为了证明 $n=3$ 时的费尔马大定理，我们需要证明以下引理。

引理 若 a 和 b 互素，且 a^2+3b^2 为一立方数，则存在整数 p 和 q，使

$$a+b\sqrt{-3}=(p+q\sqrt{-3})^3。$$

为了证明此引理，我们非常自然地对欧拉的论证作如下的进一步讨论。

i. 若 a 和 b 互素，且 a^2+3b^2 是偶数，则 $a+b\sqrt{-3}$ 可被写成

$$a+b\sqrt{-3}=(1\pm\sqrt{-3})(u+v\sqrt{-3}),$$

其中 u 和 v 为整数，符号适当选取。

因为 a^2+3b^2 是偶数，a 和 b 必有相同的奇偶性，又因为它们互素，故都必是奇数，所以都具有形式 $4n\pm1$，且 $a+b$ 或 $a-b$ 必能被 4 整除。若 $a+b$ 被 4 整除，则方程

$$\begin{aligned}4\cdot(a^2+3b^2)&=(1^2+3\cdot1^2)(a^2+3b^2)\\&=(a-3b)^2+3(a+b)^2\end{aligned}$$

能被 4 整除。设 $(a^2+3b^2)/4=u^2+3v^2, u=(a-3b)/4$，$v=a+b/4$。利用 u 和 v，再注意到其方程等价于

$$u+v\sqrt{-3}=(a+b\sqrt{-3})(1+\sqrt{-3})/4,$$

即 $$(1-\sqrt{-3})(u+v\sqrt{-3})=a+b\sqrt{-3}。$$

故得所需方程。类似地，若 $a-b$ 被 4 整除，则

$$a+b\sqrt{-3}=(1-\sqrt{-3})(u+v\sqrt{-3})$$

u 和 v 适当选取，注意到 u 和 v 互素（否则 a 和 b 不互素），且 $a^2+3b^2=4(u^2+3v^2)$，可得同样结论。

ii. 若 a，b 互素，a^2+3b^2 可被一奇素数 w 整除，则 w 可被写成 $w=p^2+3q^2$（p 和 q 为正整数)，且 $a+b\sqrt{-3}$ 可被写成

$a+b\sqrt{-3}=(p\pm q\sqrt{-3})(u+v\sqrt{-3})$，

其中正负号适当选取，且 u，v 为整数。

这第一个结论正是上一节的断言⑤。正如欧拉的证明所知，$pb+aq$ 或 $pb-aq$ 可被 w 整除。若 $pb+aq$ 被 w 整除，则方程

$$w(a^2+3b^2)=(p^2+3q^2)(a^2+3b^2)$$
$$=(pa-3qb)^2+3(pb+aq)^2$$

能被 w^2 整除。将 $(a^2+3b^2)/w$ 写成 u^2+3v^2 的形式，

$u=(pa-3qb)/w, v=(pb+aq)/w$，

即 $$u+v\sqrt{-3}=(p+q\sqrt{-3})(a+b\sqrt{-3})/w。$$

等式两边同乘 $p-q\sqrt{-3}$，则有

$$(p-q\sqrt{-3})(u+v\sqrt{-3})=a+b\sqrt{-3}。$$

即得所求。类似地，若 $pb-aq$ 被 w 整除，则

$$a+b\sqrt{-3}=(p+q\sqrt{-3})(u+v\sqrt{-3}),$$

u 和 v 互素，且 $a^2+3b^2=w(u^2+3v^2)$。

iii. 假定 a，b 互素，则 $a+b\sqrt{-3}$可被写成

$$a+b\sqrt{-3}=\pm(p_1\pm q_1\sqrt{-3})(p_2\pm q_2\sqrt{-3})\cdots$$
$$(p_n\pm q_n\sqrt{-3}),$$

其中 p_i, q_i 为正整数，且 $p_i^2+3q_i^2$ 是4或一奇素数，$i=1,2,\cdots,n$。

若 a^2+3b^2 是偶数，则它可被整除。若 a^2+3b^2 不是1，则它有一个因子 w 等于4或一个奇素数，且由结论(1)，(2)，得到

$$a+b\sqrt{-3}=(p\pm q\sqrt{-3})(u+v\sqrt{-3}),$$

其中 $p^2+3q^2=w$。则 u，v 互素，从 $u+v\sqrt{-3}$中取出一个因子 $p\pm q\sqrt{-3}$，是和从 $a+b\sqrt{-3}$中取出一个因子一样，除了 $u^2+3v^2=(a^2+3b^2)/w$ 比a^2+3b^2 小，重复此过程，最终将得到

$$a+b\sqrt{-3}=(p_1\pm q_1\sqrt{-3})\cdots$$
$$(p_n\pm q_n\sqrt{-3})(u+v\sqrt{-3}),$$

其中 $u^2+3v^2=1$。则 $u=\pm1$，$v=0$，$u+v\sqrt{-3}=\pm1$，故得所求。

iv. 若 a，b 互素，则在上面 $a+b\sqrt{-3}$ 的分解中，除了符号的选取外，因子是完全确定的，且

$$(p_1^2+3q_1^2)(p_2^2+3q_2^2)\cdots(p_n^2+3q_n^2)=a^2+3b^2$$

是 a^2+3b^2 的因子为奇素数或 4 的一个分解。进而，若因子 $p+q\sqrt{-3}$ 出现，则因子 $p-q\sqrt{-3}$ 将不会出现。反之亦然。

这第一个结论所要证的就是由式子 $p^2+3q^2=w$ 可以确定 p，q 及其正负号，这里 w 是 4 或一个奇素数。当 $w=4$ 时，显然成立。若 w 是奇素数，且 a^2+3b^2 是 w 的另一种表示，则由断言 ii 知

$$a+b\sqrt{-3}=(p\pm q\sqrt{-3})(u+v\sqrt{-3}),$$

且 $w=w(u^2+3v^2)$，即

$$u^2+3v^2=1, u=\pm 1, v=0,$$

$$a+b\sqrt{-3}=\pm(p+q\sqrt{-3})。$$

故得证。至于第二个结论，只需由 $p+q\sqrt{-3}$ 和 $p-q\sqrt{-3}$ 相乘得到因子 p^2+3q^2，这与 a，b 互素矛盾。

到此，现在证明费尔马大定理 $n=3$ 的情况。欧拉的证明中所用到的引理，现在可容易地给出。其具体证明如下。

证明 假定

$$a^2+3b^2=p_1p_1\cdots p_n$$

为断言（iv）所述的因子为 4 或奇素数的一个分解。若此分解恰好含有 k 个 4 的因子，则 2^{2k} 能够整除 a^2+3b^2 的 2 的最大幂次。又因为 a^2+3b^2 是一个立方数，因此 $2k$ 和 k 是 3 的倍数。从而，在分解中的任何奇素数 w 的重数必定是 3 的倍数。这样 n 可被 3 整除，且因子 $p_1p_2\cdots p_n$ 可按 $p_{3k+1}=p_{3k+2}=p_{3k+3}$ 这种方式排列。因此，在由断言 iii 给出的 $a+b\sqrt{-3}$ 的分解中相应于这三个 w 的因子是相同的（这是因为唯一的选择是对 $p\pm q\sqrt{-3}$ 的正负号的选择，但这两个符号不能同时出现）。从这三组中各取出一个因子使它们相乘，则得到一个数 $c+d\sqrt{-3}$，使得

$$a+b\sqrt{-3}=\pm(c+d\sqrt{-3})^3。$$

因 $-(c+d\sqrt{-3})^3=(-c-d\sqrt{-3})^3$，故问题得证。（证完）

5. 关于两个平方数之和的注记

1654年，费尔马在一封信中（可以断言欧拉不知道此信），进一步提出将一个数有效地分解为两个平方数之和的问题，即给定一个素数，例如53，找一个把它分解为两个平方数之和的一般规律。

欧拉用间接证明的方法来解决这个问题，也就是由一个矛盾的假设推出结论，即若一个形为 $4n+1$ 的素数不是两个平方数之和，则可以找到一个无限递降的正整数序列。所以没有解决费尔马的找一个构造性方法问题。然而，若对欧拉的证明作仔细的研究，发现它可以被修改成一个构造性的证明。

欧拉在证明的第5步中是具有一定的构造性的：它说明，若 $4n+1$ 是素数，则数组

$$1^{2n}+2^{2n},\ 2^{2n}+3^{2n},\ \cdots,\ (4n-1)^{2n}+(4n)^{2n}$$

中至少有一个数可被 $4n+1$ 整除。在费尔马的例子中 $4n+1=53$，可以容易找到 $1+2^{26}$ 可被53整除，用同余表示，有

$$2^6=64\equiv 11\ (\mathrm{mod}53),$$
$$2^{12}\equiv 11^2=121\equiv 15\ (\mathrm{mod}53),$$
$$2^{13}\equiv 2\cdot 15=30\ (\mathrm{mod}53),$$
$$2^{26}\equiv 900\equiv -1\ (\mathrm{mod}53)。$$

因此，$2^{26}+1$ 可以被53整除。53能整除 $2^{26}+1$ 这一事实，在欧拉的证明中是用来指出53可以整除两个平方数之和。上面的计算，给出了一个具体的可以被53整除的两个平方数之和，即 $30^2+1=17\cdot 53$。欧拉在断言iv中总结出53必是两个平方数之和，因为否则将可构造一个无限递减序列。现在的问题是给出一个直接证明，而不是通过矛盾推导出来。

对于一般的情况 $p=4n+1$，可以利用欧拉的方法去找到 $a^2+b^2=kp$，其中 a，b 互素，$k<p$。问题是怎样被 k 分解。欧

拉证明中的第二步指出了怎样被可以写成两个平方数之和的素数分解。容易看出，a^2+b^2 的任何素因子是 2 或具有 $4n+1$ 形式的数，所以可以通过将 k 分解成素因子，然后将每个因子写成两个平方数之和，再由它们逐个相除。例如 $k=17$ 本身是一素数，可以容易地写出 $17=1^2+4^2$，这分解过程为

$$17\cdot17\cdot53=(1^2+4^2)(30^2+1^2)$$
$$=(30\mp4)^2+(1\pm120)^2$$

（选择正负号使得这些数被 17 整除），

于是得到

$$53=\left(\frac{34}{17}\right)^2+\left(\frac{119}{17}\right)^2=2^2+7^2。$$

该方法把将素数 p 写成两个平方数之和的问题归结为写成形为 $4n+1$ 的更小素数问题，即可写成两平方数之和的 k 的素因子。这也可能是费尔马本人在如下描述他的无限递减的证明时所想到的方法：若一个形为 $4n+1$ 的素数不是两个平方数之和，则存在与此数有同样性质且比此数小的素数，同样将有第三个数，依次类推，无限递减直到 5，这是具有此性质的最小数，因比 5 小的数，将不是两平方数之和，从这里我们可以推测（由归纳到不可能）出所有具有此性质的数是两平方数之和。

然而，此方法是非常冗长的，因为它包括验证 k 是一个素数，若可能进行分解，总是一个困难的步骤，然后把因子表示为两个平方数之和。一般地，一个更好的被 k 分解的方法是乘上一个更小的数 n，具体作法如下：

正如欧拉证明的断言 iv，设方程 $a^2+b^2=kp$，现在去找一个形为两个平方数之和更小的 k 的倍数的数。设 $a=q_1k\pm c$，$b=q_2k\pm d$，注意到 k 必整除 c^2+d^2，设 $c^2+d^2=nk$，因 $|c|\leqslant\frac{1}{2}k$，$|d|\leqslant\frac{1}{2}k$，故 $n\leqslant\frac{1}{2}k$。从而

$$nkkp=(c^2+d^2)(a^2+b^2)=(ca\mp db)^2+(cb\pm da)^2。$$

目的在于用 k^2 除此方程。前面在证明 $cb+da$ 或 $cb-da$ 被 k 整除所用的方法不能用在这里，因为 k 不一定是素数。但有一个更简单的方法，即注意到

$$cb \pm da = c(q_1k \pm d) \pm d(q_2k \pm c)$$

可被 k 整除（这里所选的正负号要使 $cd \pm dc$ 这两项能够消去），所以整个方程能被 k^2 整除，得到 np 是两个平方数之和。若 $n=1$，则 p 即为两平方数之和，否则可重复此过程，得到一数 $m \leqslant \frac{1}{2}n$，使得 mp 是两平方数之和。重复此过程，最后必得 p 为两个平方数之和。

在例子 $30^2+1^2=17\cdot53$ 中，$c=-4$，$d=1$，$4^2+1^2=17$，此过程如前所述。再例如 $p=229$。

第一步是计算 $2^{114}\bmod 229$，有

$$2^8=256\equiv27 \pmod{229};$$
$$2^{16}\equiv27^2=929\equiv42 \pmod{229};$$
$$2^{96}\equiv(-68)(44)=-2992\equiv-15 \pmod{229};$$
$$2^{112}\equiv(-15)(42)=-630\equiv57 \pmod{229};$$
$$2^{114}\equiv4057\equiv-1 \pmod{229}。$$

所以　　229 整除 $(2^{57})^2+1^2$。

第二步是计算 $2^{57}\bmod 229$。有

$$2^{48}\equiv(-68)(42)=-2856\equiv-108;$$
$$2^{56}\equiv(-108)(27)=-2916\equiv-168\equiv61;$$
$$2^{57}\equiv122 \pmod{229}。$$

直接计算知　$122^2+1=65\cdot229$，由此开始作题，第一步 $c=122-2\cdot65=-8$，$d=1$，由此 $8^2+1^2=65$。则 $65\cdot65\cdot229=(8^2+1^2)\cdot(122^2+1^2)=(976\mp1)^2+(8\pm122)^2$，由此得

$$229=\left(\frac{975}{65}\right)^2+\left(\frac{130}{65}\right)^2=15^2+2^2。$$

类似上述过程，我们可以求出形为 $3n+1$ 的素数表示为形为 a^2+3b^2 的数。第一步是计算 $2^n\bmod p$。若 2^n 是 p 的倍数加

1，则 p 整除 2^n-1，且 $3^n \bmod p$ 必被计算。最后，必将得到一个整数 c，使得 c^n 不是 p 的倍数加 1，而 $(c-1)^n$ 是 p 的倍数加 1。则因为 p 整除 $c^{3n}-(c-1)^{3n}$，p 必整除 $c^{2n}+c^n(c-1)^n+(c-1)^{2n}$，此数是有形式 a^2+3b^2。于是 $a^2+3b^2=kp$，$k\leqslant p$。若 a，b 都是偶数，则 2 的因子可被消去。若 a，b 都为奇数，则 4 整除 a^2+3b^2，消去 4 的技巧为

$$\frac{a^2+3b^2}{4}=\left(\frac{a\mp 3b}{4}\right)^2+3\left(\frac{a\pm b}{4}\right)^2,$$

其中正负号的选取使得 4 整除 $a\pm b$。这样把问题归结到

$$a^2+3b^2=kp,$$

其中 a，b 为一奇一偶。若 $k=1$，则问题解决。否则，a 能被简化到 $c \bmod k$，b 被简化到 d，求 $c^2+d^2=nk$，$n<k$（因 k 为奇数）。则

$$nkkp=(ac\mp 3bd)^2+3(ad\pm bc)^2,$$

当符号适当选取时，可被整除，给出

$$np=e^2+3f^2。$$

若 $n\neq 1$，则可再被简化，直到

$$p=g^2+3h^2。$$

例如，考虑 $p=67$。第一步是计算 $2^{22}\bmod 27$，这是一个非常简单的计算。

$$2^6=64\equiv -3,$$
$$2^{12}\equiv 9,$$
$$2^{18}\equiv -27,$$
$$2^{19}\equiv -54\equiv 13,$$
$$2^{21}\equiv 52\equiv -15,$$
$$2^{22}\equiv -30。$$

于是 $2^{22}-1$ 不能被 67 整除，且 $2^{44}+2^{22}+1$ 必能被 67 整除。这是

$$3\left(\frac{1}{2}2^{22}\right)^2+\left(\frac{1}{2}2^{22}+1\right)^2=3(2^{21})^2+(2^{21}+1)^2。$$

因为 $2^{21}\equiv -15$，这意味着 $3\cdot(-15)^2+(-14)^2$ 必能被 67 整除。直接计算得

$$3\cdot 15^2+14^2=871=13\cdot 67。$$

简化 15 和 14mod13，得到

$$3\cdot 2^2+1^2\equiv 0\pmod{13}。$$

事实上 $3\cdot 2^2+1^2=13$，所以

$$\begin{aligned}13\cdot 13\cdot 67&=(1^2+3\cdot 2^2)(14^2+3\cdot 15^2)\\&=(14\mp 3\cdot 30)^2+3(15\pm 28)^2。\end{aligned}$$

于是

$$67=\left(\frac{104}{13}\right)^2+3\left(-\frac{13}{13}\right)^2=8^2+3\cdot 1^2。$$

6. 欧拉证明的基本思路

由上述可看出，欧拉证明费尔马大定理 $n=3$ 的情形在导出更小解的问题时，归结到如下命题：

如果 p 和 q 是互素的整数，且 p^2+3q^2 是一个立方数，则必定存在正整数 a 和 b，使得

$$p=a^3-9ab^2,\quad q=3a^2b-3b^2。$$

欧拉是怎样构思他的证明的呢？他经过深入思考，发现数 $a+b\sqrt{-3}$很有用，如果把表达式 $(a+b\sqrt{-3})^3$ 展开，命题的结论就是

$$(a+b\sqrt{-3})^3=p+q\sqrt{-3}。$$

于是命题假设 $p+3q^2$ 是一个立方。而
$p^2+3q^2=(p+q\sqrt{-3})(p-q\sqrt{-3})$，这样一来原命题可转化为如下命题：

如果 $(p+q\sqrt{-3})(p-q\sqrt{-3})$ 是一立方数，则 $p+q\sqrt{-3}$ 必定是一立方数，即

$$p+q\sqrt{-3}=(a+b\sqrt{-3})^3。$$

因此，欧拉看出引进数 $a+b\sqrt{-3}$的重要性。

欧拉对这类数 $a+b\sqrt{-3}$进行加减乘运算之后，还是该类型的数。这些性质与整数的性质十分类似。正由于这种相似性，就引导欧拉应用类比法把整数的一个性质应用到数 $a+b\sqrt{-3}$上来。由整数可唯一分解成为素数这一性质可推出整数有如下性质：互素的整数的乘积是一个立方只有当每个整数是立方才行。欧拉首先证明了，如果 p，q 互素，则 $p+q\sqrt{-3}$与 $p-q\sqrt{-3}$也互素。于是把整数的性质推广，就得到只有当每个数本身是立方时，形如 $a+b\sqrt{-3}$互素的数的乘积是立方。因此，假定 $(p+q\sqrt{-3})(p-q\sqrt{-3})$ 是立方就推出 $p+q\sqrt{-3}$ 是立方，从而命题得证。

尽管命题的结论是正确的，但是类比论证不能构成严格的逻辑论证方法。这种方法只能提供论证的方向，不能作为数学证明中的逻辑链条。它是数学家对这个问题的直觉思维的一个具体过程，但这种过程不能成为定理的一个严格证明步骤。我们通过数学直觉所获得对问题的洞察，还需要经过严格的逻辑验证。事实上，直觉和逻辑是数学创造的两个翅膀。我们学习数学不仅要学习数学家严格的推理，还要学习数学家头脑里对于定理的结论是怎样想出来的，问题是怎样提出来的，解决问题的思路一步一步又是如何进行的。而数学家欧拉的著作正体现了这两点。也正由于此，欧拉的著作才成为启发几代数学家灵感的大源泉。

欧拉用类比代替论证，这是错误的。当然，欧拉对这一错误，不一定不知道，只是由于他多产，而忽略了每一个细节，或者由于他坚信结论的正确性，就没有再去修改具体的论证细节。但这并不影响欧拉这位著名数学家的形象。错误也可以引导出正确的东西。他通过类比整数的因子分解而把形如 $a+b\sqrt{-3}$的数分解成素因子的思想，这表现出他的巨大想象力。狄利克雷（Dirichlet, Lejeune, Peter Gustar, 1805—1859）和勒让德（Legendre, Adrien-Marie, 1752—1833）对于第二个奇素数的费尔马大定理的证明所用的方法正是从欧拉证明 $n=3$ 时所用的方

法开拓出去，而相应欧拉的关键等式 $p+q\sqrt{-3}=(a+b\sqrt{-3})^3$ 的是等式 $p+q\sqrt{5}=(a+b\sqrt{5})^5$。

欧拉证明中的缺陷就在于由整数唯一因子分解所推出的性质对于数系 R 不一定成立的事实。也正是这一点，成为后来数学家深入研究费尔马大定理的一个中心课题，由此引出正则素数的概念，从而对于解决费尔马大定理的证明推进了一大步。

欧拉在证明：$4n+1$ 的每一个素数是两个平方数之和后，几乎用完全相同的方法证明了如下断言：$3n+1$ 的每个素数都具有形式 a^2+3b^2。这也是他的类比思想的具体体现。

在上述证明过程中，我们看到费尔马所创的无穷递降法得到了多次应用，由此看来这一方法有一定的普遍性。为此，我们在这里作一简单概括。费尔马无穷递降法其证明的逻辑步骤如下：

(i) 若一命题 $p(n)$ 对若干正整数 n 为真，则在此诸 n 中必有一最小者。

(ii) 若 $p(n)$ 真，则有一正整数 $n'<n$，使 $p(n')$ 亦真。

若此二步已证，则命题 $p(n)$ 决不真。

（三）关于 $n=3$ 的一个初等证明

这一节我们给费尔马大定理 $n=3$ 情形的一个初等证明。我们先把这个问题分解成几个基本命题，当这几个基本命题证明之后，就可推出费尔马大定理 $n=3$ 情形。为此先给几个引理。

引理 2.3.1 若 p 为素数，$(a,\ p)=1$，则

$$a^{p-1}\equiv 1 \pmod{p}。$$

这个引理称为费尔马小定理。为了证明该引理，我们先证下面的命题。

命题 设 p 是素数，而 h_1，$h_2\cdots$，h_a 都是整数，其中 a 为正整数，则 $(h_1+h_2+\cdots+h_a)^p\equiv h_1^p+h_2^p+\cdots+h_a^p \pmod{p}$。

证明 对 a 应用归纳法。

假设对于 $a-1$（$a>1$）时，命题成立，今证 a 亦成立。事实上，由于

$$(h_1+h_2+\cdots+h_a)^p=(h_1+h_2+\cdots+h_{a-1})^p+h_a^p+pN,$$

其中 N 为某一整数，从而

$$(h_1+h_2+\cdots+h_a)^p\equiv(h_1+h_2+\cdots+h_{a-1})^p+h_a^p(\bmod p)。$$

由归纳假设得

$$(h_1+h_2+\cdots+h_a)^p\equiv(h_1+h_2+\cdots+h_{a-1})^p+h_a^p$$
$$\equiv h_1^p+h_2^p+\cdots+h_a^p(\bmod p)。\text{(证完)}$$

现在我们来证明引理 2.3.1。

令命题中 $h_i=1$，$i=1$，2，…，a，则

$$\underbrace{(1+1+\cdots+1)^p}_{a\text{个}}=\underbrace{1^p+1^p+\cdots+1^p}_{a\text{个}}\ (\bmod p),$$

即

$$a^p\equiv a(\bmod p),$$

由此推出

$$a^{p-1}\equiv 1(\bmod p)。$$

引理 2.3.2 若不定方程

$$x^3+y^3+z^3=0,(x,y)=1$$

有整数解，则必有

$$xyz\equiv 0(\bmod 3)。$$

证明 由引理 2.3.1，有

$$x^3\equiv x(\bmod 3),$$
$$y^3\equiv y(\bmod 3),$$
$$z^3\equiv z(\bmod 3)。$$

由此推出

$$x+y+z\equiv 0(\bmod 3),$$

$$(x+y+z)^3 \equiv 0 \pmod{27},$$
$$x(x+y)+z(x+y)(x+y+z) \equiv 0 \pmod{9},$$
$$xy(x+y) \equiv 0 \pmod{3}。$$

再由 $x+y \equiv -z \pmod{3}$

推出

$$-xy(x+y) \equiv xyz \pmod{3}。$$

于是推得

$$xyz \equiv 0 \pmod{3}。$$

引理 2.3.3 若不定方程

$$x^3+y^3+z^3=0,\ (x,\ y)=1,\ xyz \neq 0 \tag{1}$$

有整数解，则存在整数 α，β，使得不定方程组

$$\begin{cases} x+y=3^{3n-1}\alpha^3,\ 3 \nmid \alpha,\ 3 \nmid \beta \\ x^2-xy+y^2=3\beta^3,\ (x,\ y)=(\alpha,\ \beta)=1 \end{cases} \tag{2}$$

也有解，反之亦然。

证明 若（1）的解为 x，y，z，则由引理 2.3.2 知，$3 \mid xyz$，不妨设 $3 \mid z$。因为 $(x,\ y)=1$，故知 $(x,\ y,\ z)=1$，所以 $3 \nmid x$，$3 \nmid y$，于是可令 $z=3^n z_0$，$3 \nmid z_0$，其中 n 为正整数。所以（1）可改写为

$$(x+y)(x^2-xy+y^2)=-3^{3n}z_0^3。$$

现在我们来证明

$$(x+y, x^2-xy+y^2)=3。$$

可从下面三种情况来考虑。

(i) $x+y$ 与 x^2-xy+y^2 没有与 3 互素的公约数。如果不然，则有 $x+y \equiv 0 \pmod{p}$，$x^2-xy+y^2 \equiv 0 \pmod{p}$，$3 \nmid p$，$p$ 为素数。由 $x^2-xy+y^2=(x+y)^2-3xy \equiv 0 \pmod{p}$，得 $xy \equiv 0 \pmod{p}$。但 $x \equiv -y \pmod{p}$，于是有 $p \mid x$，$p \mid y$，这与 $(x,\ y)=1$ 矛盾。故知 $x+y$ 与 x^2-xy+y^2 没有与 3 互素的公约数。

(ii) $x+y \equiv 0 \pmod{3}$，$x^2-xy+y^2 \equiv 0 \pmod{3}$。事实上，由

$3 \mid z$，推出 $3 \mid x^3 + y^3$，于是

$$x^3 + y^3 = (x + y)^3 - 3xy(x + y) \equiv 0(\bmod 3)$$，故知 $(x + y)^3 \equiv 0(\bmod 3)$，所以 $x + y \equiv 0(\bmod 3)$，从而有

$$x^2 - xy + y^2 = (x + y)^2 - 3xy \equiv 0(\bmod 3)。$$

(iii) $x^2 - xy + y^2 \not\equiv 0(\bmod 3^2)$。

如若不然，即有 $x^2 - xy + y^2 \equiv 0\ (\bmod 3^2)$，由 (ii) 知，$3xy \equiv 0\ (\bmod 3^2)$，于是有 $xy \equiv 0(\bmod 3)$，这表明 $3 \mid x$ 或 $3 \mid y$，而由 $3 \mid z$，则知 $3 \mid x$，$3 \mid y$，这与 $(x, y) = 1$ 矛盾。

由(i)，(ii)，(iii)就可推知

$$(x + y, x^2 - xy + y^2) = 3。$$

于是可令 $z_0 = -\alpha\beta$，$(\alpha, \beta) = 1$，而有 $3 \nmid \alpha$，$3 \nmid \beta$，则得

$$x + y = 3^{3n-1}\alpha^3,$$

$$x^2 - xy + y^2 = 3\beta^3。$$

这说明方程组 (2) 有解。

反之，若 (2) 有解，亦可推出 (1) 有解。

引理 2.3.4 当 n 为不被 3 整除的正奇数时，不定方程

$$x^2 + 3y^2 = n, \quad (x, y) = 1, \quad x > 0, \quad y > 0$$

的解数等于二次同余方程

$$z^2 + 3 \equiv 0(\bmod n)$$

的解数的一半。

证明 $x^2 + 3y^2 = n$ 的解的集合记为

$$N(x, y) = \{x, y \mid x^2 + 3y^2 = n,\ (x, y) = 1, x > 0, y > 0\}。$$

又令

$$N(z_0, z_1) = \{z_0, z_1 \mid z^2 + 3 \equiv 0(\bmod n),\ z_0 \equiv -z_1(\bmod n), n > z_0 > 0, n > z_1 > 0\}。$$

将 $N(z_0, z_1)$ 简记为 $N(z)$，其中 z 表示 z_0 或 z_1。

作映射 $\varnothing$：$x \equiv zy(\bmod n)$，x，$y \in N(x, y)$，$z \in N(z)$。下面证 $\varnothing$ 是 $N(x, y)$ 到 $N(z)$ 上的一一对应，则命题得证。为

此，分下面三种情况讨论之。

(i) $\forall x, y \in N(x, y)$，$x^2+3y^2=n$。于是 $x^2 \equiv -3y^2 \pmod{n}$。由 $x=zy \pmod{n}$，推知 $x^2 \equiv z^2y^2 \pmod{n}$，故有 $z^2y^2 \equiv -3y^2 \pmod{n}$。

因为 $(x, y)=1$，故有 $(y, n)=1$。所以，$z^2+3 \equiv 0 \pmod{n}$。

另一方面，$x \equiv z^{(1)}y \pmod{n}$，$x \equiv z^{(2)}y \pmod{n}$，由此知 $z^{(1)}y \equiv z^{(2)}y \pmod{n}$，$(y, n)=1$，
故有 $z^{(1)}=z^{(2)} \pmod{n}$。而 $z^{(1)}>0$，$z^{(2)}>0$，所以 $z^{(1)}=z^{(2)}$，于是有
$(z_0^{(1)}, z_1^{(1)})=(z_0^{(2)}, z_1^{(2)})$，故 ϕ 为 $N(x,y)$ 到 $N(z)$ 内的映射。

(ii) $\forall z_0, z_1 \in N(z)$，则必有有理数 $\frac{p_0}{q_0}$，$\frac{p_1}{q_1}$，使得

$$\frac{z_1}{n}=\frac{p_1}{q_1}+\frac{\theta_1}{q_1\sqrt{n}}, \frac{z_0}{n}=\frac{p_0}{q_0}+\frac{\theta_0}{q_0\sqrt{n}}。$$

其中 $(p_i, q_i)=1, \theta_1=-\theta_0, |\theta_i|<1, 0<q_i \leqslant \sqrt{n}, i=0,1$。

事实上，对任意实数 ξ，$\eta>1$，存在有理数 $\frac{p_1}{q_1}$，$(p_1, q_1)=1$，使得

$$\left|\xi-\frac{p_1}{q_1}\right|<\frac{1}{q_1\eta}, \quad 0<q_1 \leqslant \eta。$$

于是，存在 $|\theta_1|<1$，有

$$\left|\xi-\frac{p_1}{q_1}\right|=\frac{|\theta_1|}{q_1\eta}。$$

适当选取 θ_1 的符号，且令 $\xi=\frac{z_1}{n}$，$\eta=\sqrt{n}>1$，得

$$\frac{z_1}{n}=\frac{p_1}{q_1}+\frac{\theta_1}{q_1\sqrt{n}}, \quad 0<q_1 \leqslant \sqrt{n}, \quad |\theta_1|<1。 \tag{3}$$

令 $q_1-p_1=p_0$，$q_1=q_0$，$(p_0, q_0)=1$，$\theta_0=-\theta_1$，于是得

$$\frac{z_0}{n}=\frac{p_0}{q_0}+\frac{\theta_0}{q_0\sqrt{n}}, 0<q_0 \leqslant \sqrt{n}, |\theta_0|<1。 \tag{4}$$

由(3)，(4)两式，有

$$q_i z_i = p_i n + \theta_i \sqrt{n}, i = 0,1。$$

设 $r_i=\sqrt{n}\theta_i$，$0<|r_i|<\sqrt{n}$，r_i 显然为整数，则有

$$q_i z_i \equiv r_i (\bmod n), i = 0,1 。$$

由 $z_i^2+3\equiv 0 \ (\bmod n)$

和

$$q_i^2 z_i^2 \equiv r_i^2 \ (\bmod n)$$

有

$$r_i^2+3q_i^2\equiv 0 \ (\bmod n), \quad i=0, \ 1。$$

因为 $0<r_i^2+3q_i^2<n+3n=4n$，

故可分下面三种情况讨论之。

1° $r_i^2+3q_i^2=n$，$i=0$，1。

当 $\theta_1>0$ 时，$r_1>0$，取 $x=r_1$，$y=q_1$。

故 r_1，$q_1\in N(x, y)$，且 $r_1\equiv z_1 q_1 (\bmod n)$；

当 $\theta_1<0$ 时，$\theta_0=-\theta_1>0$，可取 $x=r_0$，$y=q_0$，

且 $r_0\equiv z_0 q_0 \ (\bmod n)$，故也有 r_0，$q_0\in N(x, y)$。

2° $r_i^2+3q_i^2=2n$，$i=0$，1。

当 r_i，q_i 同为奇数，或同为偶数，则 n 必为偶数，与题设矛盾。因此，这种情形是不可能出现的。

3° $r_i^2+3q_i^2=3n$，$i=0$，1。

上式等号两边同除以 3，得

$$q_i^2 + 3\left(\frac{r_i}{3}\right)^2 = n, i = 0,1。$$

当 $\theta_1>0$ 时，$\theta_0<0$，$r_0<0$，令 $x=q_0$，$y=-\frac{r_0}{3}$。此时，因 $r_0 \equiv z_0 q_0 \ (\bmod n)$，推知 $r_0 z_0=z_0^2 q_0 (\bmod n)$，于是 $r_0 z_0\equiv -3q_0 \ (\bmod n)$，进而推得

$$q_0 \equiv z_0\left(-\frac{r_0}{3}\right)(\bmod n),$$

故 q_0，$-\frac{r_0}{3}\in N$（x，y）；当 $\theta_1<0$ 时，$-r_1>0$，令 $x=q_1$，$y=-\frac{r_1}{3}$，同样，有 $q_1\equiv z_1\left(-\frac{r_1}{3}\right)$（$\mathrm{mod}\,n$），故 q_1，$-\frac{r_1}{3}\in N$（x，y）。

由 1°，2°，3°得知，ϕ 是满射的。

(iii) 对任意的 x_1,y_1、$x_2,y_2\in N(x,y)$，且 $(x_1,y_1)\neq(x_2,y_2)$。由此知 $x_1\neq x_2$ 或 $y_1\neq y_2$。

由 (i) 必有 $(z_0^{(1)},\ z_1^{(1)})$，$(z_0^{(2)},\ z_1^{(2)})\in N(z)$，且满足

$$x_1\equiv z_1^{(1)}y_1\ (\mathrm{mod}\,n),\quad x_2\equiv z_1^{(2)}y_2\ (\mathrm{mod}\,n)。$$

这时，必有

$$(z_0^{(1)},z_1^{(1)})\neq(z_0^{(2)},z_1^{(2)}),\text{且}$$

$$z_1^{(1)}\not\equiv -z_1^{(2)}(\mathrm{mod}\,n)。$$

事实上，若 $z_1^{(1)}\equiv -z_1^{(2)}(\mathrm{mod}\,n)$，此时有

$$x_1\equiv z_1^{(1)}y_1\ (\mathrm{mod}\,n),\quad x_2\equiv -z_1^{(1)}y_2(\mathrm{mod}\,n),$$

也就有

$$x_1y_2+x_2y_1\equiv 0(\mathrm{mod}\,n)。$$

因为 $0<(x_1y_2+x_2y_1)^2\leqslant(x_1^2+y_1^2)(x_2^2+y_2^2)<$

$$<(x_1^2+3y_1^2)(x_2^2+3y_2^2)=n^2,$$

所以　$0<x_1y_2+x_2y_1<n$。

这与　$x_1y_2+x_2y_1\equiv 0$（$\mathrm{mod}\,n$）矛盾，故

$$z_1^{(1)}\not\equiv -z_1^{(2)}\ (\mathrm{mod}\,n)。$$

又若 $(z_0^{(1)},z_1^{(1)})=(z_0^{(2)},z_1^{(2)})$，就有 $z_1^{(1)}=z_1^{(2)}$，由此可知

$$x_1y_2\equiv x_2y_1(\mathrm{mod}\,n)。$$

再由　$x_i^2+3y_i^2=n$，$i=1$，2，推出 $x_i<\sqrt{n}$，$y_i<\sqrt{n}$，$i=1$，2。此时有 $x_1y_2<n$，$x_2y_1<n$，于是有 $x_1y_2=x_2y_1$。又由 $(x_i,y_i)=1$，$i=1,2$，推出 $x_1\mid x_2$，$x_2\mid x_1$，于是有 $x_1=x_2$，$y_1=y_2$，这与 $x_1\neq x_2$ 或 $y_1\neq y_2$ 矛盾。故知 ϕ 是单射的。

由(i)，(ii)，(iii)可知 ϕ 是一一对应，而 $N(x,y)$，N

(z)均为有限集，从而知 $N(x, y)$ 中元的个数 $=N(z)$中元的个数。于是命题得证。(证完)

引理 2.3.5 设 n 为不被 3 整除的正奇数，则同余方程

$$z^2+3\equiv 0(\bmod n)$$

与同余方程

$$z^2+3\equiv 0(\bmod n^3)$$

的解数相等。

证明 分两种情形证明。

(i) 若 $n=p$ 为大于 3 的素数，则

$$z^2+3\equiv 0(\bmod p^k), k>0$$

的解数为

$$1+\left(\frac{-3}{p}\right),$$

其中$\left(\frac{-3}{p}\right)$为 Legender 符号。故当 n 为大于 3 的素数时。引理成立。

(ii) 当 n 为不被 3 整除的正奇数时，可令 $n=p_1^{k_1}p_2^{k_2}\cdots p_l^{k_l}$。若同余方程

$$z^2+3\equiv 0(\bmod p^{2k_i}), i=1,2,\cdots,l$$

的解数为 T_i，则同余方程

$$z^2+3\equiv 0(\bmod n^3)$$

的解数为

$$T=T_1\cdot T_2\cdots T_l。$$

由 (i) 知，同余方程

$$z^2+3\equiv 0(\bmod n)$$

的解数也等于 T，于是引理证毕。

推论 若 n 为不被 3 整除的正奇数时，则不定方程

$$x^2+3y^2=n,(x,y)=1$$

与不定方程

$$x^2+3y^2=n^3,(x,y)=1$$

的解数的个数相等。

引理 2.3.6 不定方程

$$x^2+3y^2=z^3,(x,y)=1 \tag{5}$$

的一切整数解都可表为

$$\begin{cases} x=u^3-9uv^2, \\ y=3u^2v-3v^3,(u,v)=1,3\nmid u, \\ z=u^3+3v^2,u,v\text{ 为一奇一偶。} \end{cases} \tag{6}$$

证明 N_1 记为（5）的解集合，N_2 记为（6）的解集合。分下面两种情况证明。

(i) 对任意 x，y，$z\in N_2$，易验证

$$(u^3-9uv^2)^2+3(3u^2v-3v^3)^2=(u^2+3v^2)^3,$$

且 $(u,v)=1,3\nmid u$，u,v 为一奇一偶。由此推出 $(u,3v)=1,(u,u+v)=1,(u,u-v)=1,(u+3v),(u-3v)$ 分别与 $u+v$，$u-v$ 互素。

由(6)推出 $(x,y)=1$，所以 $N_2\subset N_1$。

(ii) 对于任意 x，y，$z\in N_1$，由 $(x,y)=1$ 推出 z 为不被 3 整除的正奇数。事实上，若 $3\mid z$，推出 $3\mid x$，$3\mid y$，矛盾。又若 z 为偶数，推知 x，y 均为奇数。令 $x=2k+1$，$y=2m+1$。代入(5)中可算得(5)式两边一奇一偶，这是不可能的。故 z 为不被 3 整除的正奇数。

所以对于上述的 z 由推论知 $u^2+3v^2=z$ 与 $x^2+3y^2=z^3$ 的解数相等，并且满足 $3\nmid u,(u,v)=1$，u,v 为一奇一偶，而且由(6)中 $u^2+3v^2=z$ 的每一满足条件的 u,v 确定(5)的一组解。如果能证明对某一固定的 z，由 $u^2+3v^2=z$ 所确定的两组不同的解 $(u_1,v_1),(u_2,v_2)$，则由(6)确定的 (x_1,y_1) 与 (x_2,y_2) 也不一样。便有 $N_1\subset N_2$。

我们再证：若 $(u_1,v_1)\neq(u_2,v_2)$，则 $(x_1,y_1)\neq(x_2,y_2)$。事实上，如若不然，即有 $(x_1,y_1)=(x_2,y_2)$，可知 $x_1=x_2,y_1=y_2$。

由

$$\frac{1}{3}xy = uv(u^2-9v^2)(u^2-v^2)$$
$$= uv(u^2+3v^2)^2-16u^3v^3$$
$$= uvz^2-16u^3v^3$$

可推得

$$u_1v_1z^2-16u_1^3v_1^3=u_2v_2z^2-16u_2^3v_2^3,$$

将上式移项分解，得

$$(u_1v_1-u_2v_2)[z^2-16(u_1^2v_1^2+u_1v_1u_2v_2+u_2^2v_2^2)] = 0。$$

因为 z 为正奇数，故有 $u_1v_1=u_2v_2$。由

$$u_1^2+3v_1^2=z, u_2^2+3v_2^2=z$$

有 $$u_1^2+3v_1^2=u_2^2+3v_2^2,$$

与等式 $2\sqrt{3}u_1v_1=2\sqrt{3}u_2v_2$ 相加，得

$$u_1^2+3v_1^2+2\sqrt{3}u_1v_1=u_2^2+3v_2^2+2\sqrt{3}u_2v_2,$$

即

$$(u_1+\sqrt{3}v_1)^2=(u_2+\sqrt{3}v_2)^2。$$

两边开平方得

$$u_1+\sqrt{3}v_1=\pm(u_2+\sqrt{3}v_2),$$

于是有

$$u_1=\pm u_2,\quad v_1=\pm v_2。$$

再由(6)中，$x_1=u_1(u_1^2-9v_1^2)$，

$$x_2=u_2(u_2^2-9v_2^2),$$

推知 $u_1=u_2$，$v_1=v_2$。这与$(u_1,v_1)\neq(u_2,v_2)$矛盾。

由(i)，(ii)知有 $N_1=N_2$。

引理 2.3.7 不定方程

$$x^2+3y^2=3z^3,\ (x,\ y)=1$$

的整数解可表为

$$\begin{cases} x = 9u^2v - 9v^3, \\ y = u^3 - 9uv^2, (u, v) = 1, 3 \nmid u, \\ z = u^2 + 3v^2, u, v \text{ 为一奇一偶。} \end{cases}$$

证明 若 $x^2+3y^2=3z^3$ 有解，则 $3 \mid x$，将方程变形为

$$y^2 + 3\left(\frac{x}{3}\right)^2 = z^3, \left(y, \frac{x}{3}\right) = 1。$$

由引理 2.3.6 可得

$$\frac{x}{3} = 3u^2v - 3v^3,$$

$$y = u^3 - 9uv^2, \ (u, v) = 1, \ 3 \nmid u,$$

$$z = u^2 + 3v^2, \ u, v \text{ 为一奇一偶。}$$

再变形，就得所求之解。

引理 2.3.8 不定方程

$$x^2 - xy + y^2 = 3z^3, (x, y) = 1$$

的一切整数解可表为

$$\begin{cases} x = u^3 + 9u^2v - 9uv^2 - 9v^2, \\ y = 2u^3 - 18uv^2, \\ z = u^2 + 3v^2, \end{cases} \tag{7}$$

或

$$\begin{cases} x = u^3 + 9u^2v - 9uv^2 - 9v^3, \\ y = -u^3 + 9u^2v + 9uv^2 - 9v^3, \\ z = u^2 + 3v^2, \end{cases} \tag{8}$$

其中 $(u, v) = 1$，$3 \nmid u$，u，v 为一奇一偶。

证明 易验证(7)，(8)为不定方程

$$x^2 - xy + y^2 = 3z^3, (x, y) = 1$$

的解。分下面两种情形讨论。

(i) 当 $2 \mid xy$ 时，不妨假定 $2 \mid y$，则原不定方程可改写为

$$\left(x - \frac{y}{2}\right)^2 + 3\left(\frac{y}{2}\right)^2 = 3z^3。$$

由 $(x,y)=1$，推知 $\left(x-\frac{y}{2},\frac{y}{2}\right)=1$，于是由引理 2.3.7 推知

$$x-\frac{y}{2}=9u^2v-9v^3,$$

$$\frac{y}{2}=u^3-9uv^2,\ (u,\ v)=1,\ 3\nmid u$$

$$z=u^2+3v^2,\ u,\ v\text{ 为一奇一偶。}$$

由此解出 x，y，可写成(7)的形式。

(ii) 当 $2\nmid xy$ 时，推知 $2\mid x+y$，$2\mid x-y$，于是原方程可改写为

$$\left(\frac{x+y}{2}\right)^2+3\left(\frac{x-y}{2}\right)^2=3z^3。$$

由 $(x,y)=1$，知 $\left(\frac{x+y}{2},\frac{x-y}{2}\right)=1$，于是

由引理 2.3.7，得

$$\begin{cases}\dfrac{x+y}{2}=9u^2v-9v^3,\\[2mm] \dfrac{x-y}{2}=u^3-9uv^2,\ (u,\ v)=1,\ 3\nmid u,\\[2mm] z=u^2+3v^2,\ u,\ v\text{ 为一奇一偶。}\end{cases}$$

由此，解出 x，y，可写成(8)的形式。

引理 2.3.9 不定方程

$$x^3+y^3+z^3=0,\ xyz\neq 0 \tag{9}$$

无整数解。

证明 不妨设 $(x,\ y,\ z)=1$，由引理 2.3.3 知，(9)可改写为

$$x^3+y^3+3^{3n}z_1^3=0,(x,y,z)=1, \tag{10}$$

其中 $xyz_1\neq 0$，$3\nmid x$，$3\nmid y$，$3\nmid z_1$，$n\geqslant 0$。

对 n 应用数学归纳法。

当 $n=0$ 时，显然定理真。

若 $n-1$ 真，下证对 n 亦真。如若不然，即(10)有整数解，

由引理2.3.3知，存在整数α，β，使得x，y满足

$$x+y=3^{3n-1}\alpha^3,\ (\alpha,\ \beta)=1,\ n\geqslant 1, \tag{11}$$
$$x^2-xy+y^2=3\beta^3,\ 3\nmid\alpha,\ \beta,$$

那么，由引理2.3.8中的(7)代入上式，得

$$x+y=3u^3+9u^2v-27uv^2-9v^2=3^{3n-1}\alpha^3,$$
$$u^3+3u^2v-9uv^2-3v^3=3^{3n-2}\alpha^3,\ (n\geqslant 1)。$$

由$3\nmid u^3$推知上式左边不被3整除，而右边为3的倍数，矛盾。

而引理2.3.8中的(8)代入(11)可得

$$2v(u+v)(u-v)=3^{3(n-1)}\alpha^3。$$

由$(u,\ v)=1$，$3\nmid u$，u，v为一奇一偶，推知$2v$，$u+v$，$u-v$两两互素，则其中必有一个能被$3^{3(n-1)}$整除，于是可令$\alpha=\alpha_1\beta_1 r_1$，且$\alpha_1$，$\beta_1$，$r_1$两两互素，则可得

(i) 若$3^{3(n-1)}|2v$，不妨设$2v=3^{3(n-1)}r_1^3$，$u+v=\alpha_1^3$，$u-v=\beta_1^3$，三式消去v，得

$$\alpha_1^3+(-\beta_1)^3+3^{3(n-1)}(-r_1)^3=0。 \tag{12}$$

(ii) 若$3^{3(n-1)}|u-v$，不妨设$u-v=3^{3(n-1)}\gamma_1^3$，$u+v=\alpha_1^3$，$2v=\beta_1^3$，消去u，v，得

$$\alpha_1^3+(-\beta_1)^3+3^{3(n-1)}(-r_1)^3=0。 \tag{13}$$

(iii) 若$3^{3(n-1)}|u+v$，不妨设$u+v=3^{3(n-1)}{r_1}^3$，$u-v=\beta_1^3$，$2v=\alpha_1^3$，消去u，v，得

$$\alpha_1^3+\beta_1^3+3^{3(n-1)}(-r_1)^3=0 \tag{14}$$

上面式(12)，(13)，(14)均与归纳假设矛盾，故对任意非负整数n，(10)无整数解，从而(9)无整数解。(证完)

(四) 从勒让德到库姆尔

1. 关于 $n=5$ 和 $n=7$，分圆整数

勒让德（Legendre，Adrien-Marie，1752—1833）和狄利克

雷于 1825 年对于第二个奇素数 5 的情形证明了费尔马大定理。所用的方法本质上是欧拉证明 $n=3$ 情形所用方法的推广。前面我们已经谈过与欧拉关键的等式 $p+q\sqrt{-3}=(a+b\sqrt{-3})^3$ 相应的等式在这里是 $p+q\sqrt{5}=(a+b\sqrt{5})^5$。为了证明 $p+q\sqrt{5}=5$ 次幂，不仅要假定 p^2-5q^2 是一个五次幂且 p 与 q 互素，而且还要假定 p 与 q 是一奇一偶，且 q 被 5 整除。这就比在证明 $n=3$ 的情形要复杂得多。狄利克雷证明这一事实是基于对于形如 x^2-5y^2 的数进行研究的结果。他的证明是仿照欧拉关于对形如 x^2+3y^2 的数研究等工作以及拉格朗日和高斯的工作进行的。

1839 年，拉梅（Grabriel Lame，1795－1870）对于第三个奇素数 7 证明了费尔马大定理。但所用的证明方法十分繁难、冗长且与 7 的个性紧密联系在一起，从这个证明中丝毫窥视不出证明 n 等于其他情形的任何曙光。研究费尔马问题必须创造新的方法，才能取得突破性的进展。

1847 年，拉梅为了证明一般的费尔马大定理，他引进了 1 的 n 次复根：即 $\alpha^n=1$，其中 α 为复数，但对于任何小于 n 的正整数 k，$\alpha^k\neq1$。

注意 1，α，α^2，…，α^{n-1}是方程 $z^n-1=0$ 的根，所以根据代数基本定理有

$$z^n-1=(z-1)(z-\alpha)(z-\alpha^2)\cdots(z-\alpha^{n-1})。$$

令 $z=\dfrac{x}{y}$，然后方程两边同乘以 y。因为只要考虑 n 为奇数的情形就够了，故有

$$x^n+y^n=(x+y)(x+\alpha y)\cdots(x+a^{n-1}y)。$$

因此，我们知道 x^n+y^n 的每个因子是形如

$$a_0+a_1a+\cdots+a_{n-1}a^{n-1}$$

的数，其中 a_0，a_1，…，a_{n-1}是整数。我们把这种形式的数叫做分圆整数。分圆整数也与数 $a+b\sqrt{-3}$ 一样构成一个数环。

特别，当 $a_1=a_2=\cdots=a_{n-1}=0$ 时，这时分圆整数就是普通的整数了。因此分圆整数的集合包含了普通整数的集合。这样一来，把原来考虑不定方程解的问题从一个小范围扩展到一个大范围，这就可能为寻找解带来更大的周旋余地。

正由于分圆整数是一个数环，因此拉梅就把整数的一些性质类比到分圆整数上来。他提出如下断言：互素的分圆整数的乘积是一个 n 次幂，只当每个分圆整数是一个 n 次幂。根据这一断言，拉梅推出不定方程 $x^n+y^n=z^n$ 不可解。而上述的断言是基于分圆整数因子分解的唯一性，然而在分圆整数的集合中，因子分解的唯一性并不成立。这就导致拉梅的证明是错误的。他的错误也正如欧拉的错误是一样的，把普遍整数因子分解唯一性定理类比到分圆整数上来。

1844 年，库姆尔（Ernst Eduard Kummer，德国人，1810—1893）证明了分圆整数唯一因子分解定理通常并不成立。当他继续研究分圆整数时，想了一种办法，可以把唯一分解这一概念加以变更，使得变更后，可以用来证明分圆整数的更有用的性质。库姆尔在分圆整数中引入了一个新概念——理想素因子。正由于这个概念的引入，使分圆整数以及像 $a+b\sqrt{-3}$的数，就可以由唯一因子分解推出其最重要的性质，从而使费尔马大定理的研究取得了前所未有的最大进展。为了较深入地了解库姆尔所取得的结果，下面我们先简单介绍一下有关代数数论的知识。

2. 代数数论基本知识

欧拉在证明费尔马大定理 $n=3$ 的情形，引进了形如 $a+b\sqrt{-3}$的数，其中 a，b 为整数。这个数所组成的集合显然包含普通整数集合，它有许多性质与整数集合相类似。因此，使我们联想到，若 a，b 为整数，那么复数 $a+bi$ 所组成的集合是否也具有与普通整数（为了区别以后所说的其他整数，称普通整数为有理整数）集合相同的性质呢？下面我们就来研究这一问题。

设 z 表示全体整数所组成的集合，若 a，$b\in z$，那么复数 $a+bi$ 叫做复整数，由复整数全体所组成的集合记为

$$z(i)=\{a+bi\mid a,b\in z\}。$$

特别，当 $b=0$ 时，$z(i)=z$。

显然，当 $a+bi\in z(i)$，$c+di\in z(i)$，则有

$$a+bi+c+di=(a+c)+(b+d)i\in z(i),$$

$$a+bi-(c+di)=(a-c)+(b-d)i\in z(i),$$

$$(a+bi)(c+di)=(ac-bd)+(ad+bc)i\in z(i)。$$

设 $Q=\{\alpha|\alpha$ 为有理数$\}$。若 α，$\beta\in Q$，则称复数 $\alpha+\beta i$ 为有理数复数。由有理数复数的全体组成的集合记为

$$Q(i)=\{\alpha+\beta i\mid \alpha,\beta\in Q\}。$$

显然，$Q(i)\supset z(i)$，$Q(i)\supset Q$。

对于复数的加法和乘法，$Q(i)$组成一个数域，通常叫$Q(i)$是i添加到Q上的代数数域。

现在我们要问在 $Q(i)$中的复整数是否具有有理整数的一些类似性质呢？为此，我们先定义 $Q(i)$中的复整数的有关概念。

设 A，$B\neq0$，且 A，$B\in Z(i)$，如果存在$C\in Z(i)$，使

$$A=BC,$$

则称 B 整除A，记为 $B|A$；否则，称 B 不整除A，记为 $B\nmid A$。

在复整数中，整除 1 的有±1，$\pm i$，这四个数叫做 $Q(i)$的单位数。

若 $A=a+bi\in z(i)$，A 的共轭数为$\bar{A}=a-bi$，则 $N(A)=A\bar{A}=a^2+b^2$ 叫做 A 的范数。显然有

$$N(A)=\begin{cases}0, A=0,\\ 1, A \text{ 是单位数},\\ a^2+b^2>1, \text{其他}。\end{cases}$$

设 ε 为$Q(i)$的单位数，A，$B\in Z(i)$，如果满足方程 $A=B\varepsilon$，则称 A 与B 相结合。例如$3-i=-i(1+3i)$，所以 $3-i$ 和 $1+3i$ 是相结合的。

设 $N(A)>1$，对任何分解式

$$A=BC$$

若有 $N(B)=1$ 或 $N(C)=1$，则称 A 是 $Q(i)$ 的素数，记为 π。可证明

定理 2.4.1 $Q(i)$中的素数是 $1+i$ 和它的相结合数，有理素数 $q(q\equiv3(\mathrm{mod}4))$和它的相结合数，有理素数 $p(p\equiv1(\mathrm{mod}4))$的因数 $a+bi$。

类似有理整数，可证明复整数唯一分解定理，即有

定理 2.4.2 设 $N(A)>1$，$A=\pi_1\cdots\pi_n=\pi'_1\cdots\pi'_m$（$n\geqslant1$，$m\geqslant1$），则有 $n=m$，且诸 π_i 以适当的次序是诸 π'_j 的相结合数。

现在我们用上面所介绍的代数数论的基本性质，求出在第一章（二）中所说的不定方程的全部正整数解，即求不定方程

$$x^2+y^2=z^2,\quad x>0,y>0,z>0,$$
$$(x,y)=1,\quad 2\mid x \tag{1}$$

的整数解。

为了求（1）的解，我们把方程变形为

$$(x+yi)(x-yi)=z^2。\tag{2}$$

设 $x+yi$，$x-yi$ 的最大公因数为 d，则 $d\mid2x$，$d\mid2yi$。因为 $2\nmid z$，$(x,y)=1$，故 $d=1$。由定理 2.4.2 和(2)，得

$$x+yi=(c+di)^2,\tag{3}$$

或

$$x+yi=i(-a+bi)^2。\tag{4}$$

而 $2\nmid y$，故(3)为不可能。由(4)得

$$x+yi=2ab+(a^2-b^2)i,$$

所以 $x=2ab$，$y=a^2-b^2$。

再由(1)得 $z=a^2+b^2$，其中 $a>b>0$，a，b 为一奇一偶，且满足$(a,b)=1$ 的任意整数。这正是我们在第一章（二）中所得的结果。由此看来，利用代数数论的知识很容易把方程（1）的

解求出来了。代数数论可以用来解决某些不定方程问题。

下面，我们把数域 $Q(i)$的概念一般化。我们知道，在 $Q(i)$中任一数 $\alpha=a+bi$ 适合多项式 $x^2-2ax+a^2+b^2$，当 $\alpha\in z(i)$ 时，α 适合一个首项系数为 1 的整系数的二次多项式。一般地，设 $n>0$，如果复数 θ 是一个系数为有理数的 n 次（有理数域上）不可约多项式的根，则称 θ 为 n 次代数数。如果 θ 为一个首项系数为 1，其余系数为有理整数的 n 次不可约多项式的根，则称 θ 为 n 次代数整数。

现在介绍一般代数数域的基本概念和基本性质。

设 θ 是一个 n 次代数整数，Q 是有理数域，$Q(\theta)$表示 θ 添加到 Q 上得到的 n 次代数数域。任一数$\alpha\in Q(\theta)$，均可写为

$\alpha=a_0+a_1\theta+\cdots+a_{n-1}\theta^{n-1}$，$a_i\in Q$，$i=0$，1，…，$n-1$。记 $\theta=\theta^{(1)}$，并设 $\theta^{(2)}$，…，$\theta^{(n)}$为 θ 所适合的不可约多项式的其他 $n-1$ 个根。

设 $\alpha^{(1)}=\alpha$，称 $\alpha^{(K)}=a_0+a_1\theta^{(K)}+\cdots+a_{n-1}\theta^{(K)^{n-1}}$（$K=2$，3，…，$n$）为 α 的共轭数，又称

$$N(\alpha)=\alpha^{(1)}\cdots\alpha^{(n)}$$

为 α 的范数。

设 α 是 $Q(\theta)$中的一个代数整数，如果 α^{-1}也是代数整数，则称 α 为$Q(\theta)$的单位元。

对于非单位元的整数 $\alpha\neq0\in Q(\theta)$，如有 $Q(\theta)$中的代数整数 β，γ，且均非单位元，使 $\alpha=\beta\gamma$，则称 α 在$Q(\theta)$中可分解；否则，称 α 是$Q(\theta)$中的素数或不可分数。

$Q(\theta)$ 中任一非单位元的整数可分解为素数的乘积。但是，在许多代数数域中素因数的唯一分解性不成立。

设 α_1，…，α_q 为$Q(\theta)$内任意 q 个代数整数，称所有形如

$$r_1\alpha_1+\cdots+r_q\alpha_q(r_1,\cdots,r_q \text{ 为 } Q(\theta) \text{ 中的整数})$$

的代数整数组成的集为由 α_1，…，α_q 生成的理想数，记为 [α_1,

$\cdots$，α_q]。设 $A=[\alpha_1,\cdots,\alpha_q]$，$B=[\beta_1,\cdots,\beta_r]$，定义

$$AB=[\alpha_1\beta_1,\cdots,\alpha_1\beta_r,\cdots,\alpha_q\beta_1,\cdots,\alpha_q\beta_r]。$$

若一理想数除了单位理想数[1]和本身以外无其他因子,则称为素理想数。只由一个代数整数 α 生成的理想数$[\alpha]$称为主理想数。

定理 2.4.3 任一不同于单位理想数 [1] 和 [0] 的理想数 A，可以分解为素理想数的乘积，且如果不计其排列的次序时分解法唯一。

设 α，β 是 $Q(\theta)$ 中的代数整数，若 $A|[\alpha-\beta]$，则称 α 和 β 对模 A 同余，记为 $\alpha\equiv\beta\ (\mathrm{mod} A)$。根据这一同余关系可以将域 $Q(\theta)$ 中的全体代数整数模 A 进行分类，其类数是有限的，记为 $N(A)$，该数叫做理想数 A 的范数。

设 A，B 是 $Q(\theta)$上的理想数,如果有 $Q(\theta)$上的主理想数 $[\alpha]$ 和 $[\beta]$，使

$$[\alpha]A=[\beta]B$$

成立，称 A 与 B 属于同一理想类，以 $A\sim B$ 记之。由此，可将 $Q(\theta)$ 上的全体理想数进行分类，称为理想数类，其类数为一有限数，以 h 表之。

定理 2.4.4 对 $Q(\theta)$上任一理想数 A，总有

$$A^h=[\beta],$$

其中 β 是 $Q(\theta)$中的一个代数整数。

设 $\eta=e^{\frac{2\pi i}{p}}$，则 η，η^2，$\cdots$，η^{p-1}是有理数域 Q 上不可约多项式

$$F(x)=x^{p-1}+x^{p-2}+\cdots+x+1$$

的全部根。显然有

$$F(x)=\prod_{k=1}^{p-1}(x-\eta^K)。$$

域 $Q(\eta)=Q(\eta^2)=\cdots=Q(\eta^{p-1})$叫$Q$ 上次数为 $p-1$ 的分圆域。

定理 2.4.5 设 $\pi=1-\eta$，则有 $Q(\eta)$中理想数

$$[p]=[\pi]^{p-1}$$

且有 $N(\pi)=p$，π 为素数。

设 $Q(\eta)$的类数为 h，如果 $p\nmid h$，则 p 叫做正则素数；如果 $p\mid h$，则 p 叫非正则素数。

3．关于正则素数

1847 年，库姆尔得到费尔马大定理成立的充分条件，即

定理 2.4.6 对于正则素数 p，

$$x^p+y^p=z^p$$

没有使 $xyz\neq0$ 的整数解。

在小于 100 的素数中，除了 37，59，67 外其余都是正则素数，这就是说，对于小于 100 的素数，除了 37，59，67 外费尔马大定理均成立。

如何判定一个素数为正则的呢？有如下方法。由下式

$$\frac{t}{e^t-1}=1-\frac{t}{2}+\sum_{m=1}^{\infty}\frac{(-1)^{m-1}Bmt^{2m}}{(2m)!}。$$

定义贝努力数$B_m(m=1,2,\cdots)$。

经计算知

$$B_1=\frac{1}{6},\ B_2=\frac{1}{30},\ B_3=\frac{1}{42},\ B_4=\frac{1}{30},$$

$$B_5=\frac{5}{66},\ B_6=\frac{691}{2730},\ B_7=\frac{7}{6},$$

$$B_8=\frac{3617}{510},\ B_9=\frac{43867}{798},\ B_{10}=\frac{174611}{330},$$

等等。

定理 2.4.7 设 $p>3$，如果奇素数 p 不整除前$\frac{1}{2}(p-3)$个贝努力数的每一个的分子，则 p 是正则素数。

通过上面的定理我们就可判定一个奇素数是否为正则素数。

100 以内的非正则数只有 3 个，在 $3\leqslant p\leqslant4001$ 的素数 p 中，有 334 个正则素数，216 个非正则素数。

通过直觉和数值计算来看，可猜想存在无穷多个正则素数。

但此断言至今未获得证明。我们知道，素数分为正则素数与非正则素数。经计算表明 60% 的素数是正则的，因此有理由相信在素数集合中多数是正则素数，少数是非正则素数。但被认为较大的子集却不能证明有无穷多元素，而对其余集倒证明了有无限多个形如 $4k+3$ 的非正则素数。

4. 其它一些结果

1857 年，库姆尔对于 100 以内的三个非正则素数 37，59，67，给出了证明，但存在一些缺点。1892 年米里曼诺夫(Dimitry Mirimanoff，1861－1945）对于 $n=37$ 时的费尔马大定理给出一个严格的证明。

对于费尔马问题

$$x^p+y^p=z^p,$$

如果加上条件 $(xyz,p)=1$，叫做费尔马大定理的第一种情形；如果加上条件 $(xyz,p)=p$，叫做费尔马大定理第二种情形。

在 $Q(\theta)$的代数整数环中考虑费尔马问题第一种情形为

$$\alpha^p+\beta^p+\gamma^p=0,(\alpha\beta\gamma,p)=1 \tag{5}$$

在 $Q(\theta)$中没有整数解。第二种情形等价变形为

$$\alpha^p+\beta^p=\varepsilon\lambda^{np}\gamma^p,(\alpha\beta\gamma,p)=1 \tag{6}$$

(其中 n 为自然数，ε 为$Q(\theta)$的单位元，$\lambda=1-\theta$)在 $Q(\theta)$中没有整数解。

设 $Q(\theta)$的类数为 h，1847 年库姆尔得到

定理 2.4.8 如果 $(h,\ p)=1$，(5)与(6)均没有解。

库姆尔的理想素数以及与它相关的某些数类现在都叫做理想数。当前理想数论已发展成为数学中一个专门分支。库姆尔引入理想素数这一概念不仅对于研究费尔马大定理具有重要作用，而且由此发展出理想数论这样一门新的数学学科。这一事实充分说明了在解决数学难题的过程中，通过理论思维的能动作用，可以产生新的数学思想，从而丰富和发展数学理论。

库姆尔是怎样引进理想素数概念的呢？他是在继承前人研究成果的基础上，充分发挥他的科学想象力开拓出来的。欧拉基于唯一因子分解定理，在形如 $a+b\sqrt{-3}$的集合中，证明了费尔马大定理 $n=3$ 的情形。尽管这一结果是对的，但证明的方法是错误的。拉梅基于唯一因子分解定理在分圆整数集合中证明了费尔马大定理，这一证明方法和在分圆整数集合因子分解唯一的结果都是错误的。错误是成功的先导，从错误中使人得到启示。无论是欧拉的证明，还是拉梅的证明，都说明唯一因子分解定理对于证明费尔马大定理是多么重要，只要唯一，就可推导出费尔马大定理成立。因此，就使人们想到，在什么情况下才能使因子分解唯一呢？这就需要把原来所考虑问题的范围缩小，从而引导出理想素数的概念来，创造出理想数论的方法。新学科的诞生是置根于已有问题研究的土壤之中。只有继承，才能开拓，一个有作为的科学开创者，必须了解所研究问题的来龙去脉，要掌握前人已取得的成果，即使是错误的，也同样具有启发性，然后发挥高度的科学想象力，这样才有可能创立科学新成果，开辟科学新领域，从而推动科学的发展。

（五）费尔马大定理研究的一些新成果

1. 考虑结论反面的必要条件

我们也可以从费尔马大定理相反的结论考虑问题，也就是说，如果不定方程有正整数解，看能推出什么结论，由此可判断方程没有正整数解的范围。

1905 年，米里曼诺夫得到如下结果：

定理 2.5.1 在第一种情形，如果

$$x^p+y^p=z^p$$

具有 $xyz\neq0$ 的整数解，则对于所有的 $-t=x/y$，y/x，y/z，

z/y，z/x，x/z，有

$$B_{2m}f_{p-2m}(t)\equiv 0(\bmod p),$$
$$m=1,2,\cdots,(p-3)/2,$$

其中 B_m 为第 m 个贝努力数，$f_m(t)=\sum_{r=0}^{p-1}r^{m-1}t^r$。这叫做库姆尔判据。

将如下形式整理变形，于 1909 年，外斐力什（A. Wieferich）证明

定理 2.5.2 在第一种情形，如果

$$x^p+y^p=z^p$$

具有 $xyz\neq 0$ 的整数解，则

$$(2^{p-1}-1)/p\equiv 0(\bmod p) \tag{1}$$

成立。

但适合上面条件的 p 相当少，当 $p<2000$ 时，只有一个数 1093，使（1）成立；当 $p<3700$ 时，只有 1093，3511 两个数，使（1）成立；1963 年，确定了当 $p<200183$ 时，也只有这两个数，使（1）成立；后来确定了，当 $p<31059000$ 时，也是只有这两个数，使（1）式成立。

上面的定理被后来的数学家作了一系列的推广，例如，1912 年，富特王格尔（Ph. Furtwängler）得到如下结论：

定理 2.5.3 在第一种情形，如果

$$x^p+y^p=z^p$$

具有 $xyz\neq 0$ 的整数解，则有

$$(3^{p-1}-1)/p\equiv 0(\bmod p)。$$

1940 年，罗塞尔（Ja. B. Rosser）得到

定理 2.5.4 在第一种情形，如果

$$x^p+y^p=z^p$$

具有 $xyz\neq 0$ 的整数解，则有

$$(m^{p-1}-1)/p\equiv 0(\bmod p), \tag{2}$$

对于 $2\leqslant m\leqslant 43$ 的所有的 m 均成立。

1941 年，罗塞尔应用这个定理，证明了当 $p<41000000$ 时，同一年，D.H.Lehmer 和 E.Lehmer 更进一步改进这一结果，证明当 $p<253749889$ 时，费尔马大定理第一种情形成立。

1934 年，日本的森岛得到

定理 2.5.5 如果（四）中（5）在 $Q(\eta)$ 中具有整数解 α，β，γ，则（2）对于 $2\leqslant m\leqslant 43$ 中所有的 m 均成立。

2. 充分条件

设 p 为奇素数，η 为 1 的本原 p 次幂根，$Q(\eta)$ 为在有理数域 Q 上添加 η 后得的分圆域，h 为 $Q(\eta)$ 的类数，设 $Q(\eta)$ 中包含的实域 $Q(\eta+\eta^{-1})$ 的类数为 h_2，则 h_2 为 h 的因子，$h_1=h/h_2$ 称为 h 的第一因子，h_2 称为 h 的第二因子。

继库姆尔之后，数学家对费尔马大定理的研究又得到了更广泛的充分条件。1929 年，美国的得克萨斯大学范迪威尔（H.S.Vandiver）得到

定理 2.5.6 如果 $(h_2, p)=1$ 和贝努力数 $B_{2mp}(m=1,2,\cdots,(p-3)/2)$ 的分子都不能用 p^3 整除，则

$$x^p+y^p=z^p \tag{3}$$

没有 $xyz\neq 0$ 的整数解。

1928—1936 年，范迪威尔，通过计算证实，对于 $2<p<619$，$x^p+y^p=z^p$ 没有正整数解。

1944 年，谢立费力基（J.L.Selfridge），尼可（C.A.Nicol），范迪威尔，证明当 $2<p<4002$ 时，（3）没有正整数解。

1975 年，约翰逊（W.Johnson）通过电子计算机用 D.H.Lehmer，E.Lehmer，H.S.Vandver 的方法，证实当 $2<p\leqslant 30000$ 时，$x^p+y^p=z^p$ 没有正整数解。1977 年，瓦格斯塔夫（Samuel S.Wagstaff）在大型计算机的帮助下，证明当 $2<p<125000$ 时，$x^p+y^p=z^p$ 无正整数解。

1929年，范迪威尔得到

定理2.5.7 如果$(h_2, p) = 1$及贝努力数$B_{2mp}(m = 1, 2, \cdots, (p-3)/2)$的分子都不能用$p^3$整除，则(四)中的(5)、(6)均没有解，但在(6)中α, β, γ限于是$Q(\eta + \eta^{-1})$的整数，并且是互素的，其中λ代以$(1-\eta)(1-\eta^{-1})$。

1934年，森岛得到

定理2.5.8 如果$(h_2, p) = 1$，则(四)中的(5)没有解，但是α, β, γ限于是$Q(\eta + \eta^{-1})$的整数。

(六) 简　　评

通过如上所述，要找一个反例来否定费尔马大定理不成立，如果这个反例存在的话，电子计算机计算结果已经表明，其数字是相当大的，它不但超过人工计算，也大大超过现有电子计算机的计算能力。由此看来，费尔马大定理是成立的。但是从数学的角度来看，能用电子计算机验证的数字再大，也不能断定费尔马大定理是正确的，只能增强对费尔马大定理相信的程度，要承认像费尔马大定理这样与自然数有关的一个数学命题是成立的，还必须从数学上给予严格的证明。

在历史上，人们从不同的角度去探索费尔马大定理的证明。费尔马、欧拉用初等数论的方法解决了$n = 3, 4$的情形。一般说来，用初等方法往往只能解决费尔马大定理的第一种情形。例如用初等方法，可得如下三个定理。

定理2.6.1 如果存在一个奇素数q，使得同余式

$$x^p + y^p + z^p \equiv 0 \pmod{q}$$

只有$q \mid xyz$的整数解x, y, z，且对任意整数K有$K^p \not\equiv p \pmod{q}$，则

$$x^p + y^p + z^p = 0, (x, y) = (x, z) = (y, z) = 1 \quad (1)$$

没有$(xyz, p) = 1$的整数解。

定理 2.6.2 设 $q=2hp+1$ 是素数，如果 $q\nmid D2h$，其中

$$D_{2h}=\begin{vmatrix}\binom{2h}{1} & \binom{2h}{2} & \cdots & \binom{2h}{2h-1} & 1\\ \binom{2h}{2} & \binom{2h}{3} & \cdots & 1 & \binom{2h}{1}\\ \cdots & & \cdots & & \cdots\\ 1 & \binom{2h}{1} & \cdots & \binom{2h}{2h-2} & \binom{2h}{2h-1}\end{vmatrix}$$

且 $p^{2h}\not\equiv 1(\bmod q)$，则方程(1) 无
$(xyz,p)=1$ 的整数解。

定理 2.6.3 设 p 是一个奇素数，当

i）$2p+1$ 是一个素数，或

ii）$4p+1$ 是一个素数时，

费尔马大定理第一种情形成立。

由定理 2.6.2，还可证明当 $8p+1$，$10p+1$，$14p+1$，$16p+1$ 之一为素数时，费尔马大定理第一种情形成立。

使费尔马大定理的研究进展速度快，取得的成果显著，是由于创造了新方法。二百年来，人们对费尔马大定理只证明了前三个奇素数成立，由于库姆尔创造了理想数论的方法，一下子使费尔马大定理成立的奇素数扩大到 100 之内。从反面考虑费尔马大定理成立，可能存在无限多个整数解，可是法尔廷斯，利用代数几何的方法证明了莫德尔猜想，由莫德尔猜想推知只能存在有限多个整数解。这种从无限一下子推进到有限，也是由于采用了现代的数学理论和方法。

在历史上，人们探求费尔马大定理的两次重大突破，都不是局限在初等数论范围以内，而是另辟蹊径，开创和利用代数数论和代数几何的方法。

费尔马大定理，由于问题简单，易于理解，再加上科学机构的宣传，因此，在过去一百年来，费尔马大定理成为业余数学爱

好者的最喜爱的题目之一。布鲁塞尔和巴黎科学院曾设奖金数次。最后一次设奖悬赏是 1908 年，数学家佛尔夫斯克尔（F. Paul Wolfsuehl）在格廷根皇家科学会以悬赏十万马克的巨款，赠给第一个证明这个定理的人。由于这几次悬赏使许多人都卷入证明费尔马大定理的行列里来。当时，证明费尔马大定理的稿件，像雪片一样飞来，但其证明都是错误的。至今仍然有许多人去探索费尔马大定理的证明，并有不少人坚信能用初等方法证明费尔马大定理，例如有人，通过如下九个引理完成对这个定理的证明。

引理 2.6.1 p 为奇素数，m 为大于 1 的整数，当 $(m-1, p)=1$ 时，则

$$\begin{aligned} m^p-1 &= (m-1)(m^{p-1}+m^{p-2}+\cdots+m+1) \\ &= (m-1)(2pB_1+1); \end{aligned}$$

当 $(m-1, p)=p$ 时，则

$$\begin{aligned} m^p-1 &= (m-1)(m^{p-1}+m^{p-2}+\cdots+m+1) \\ &= (m-1)p(2pB_2+1)。 \end{aligned}$$

设 q 为奇素数，$q\mid 2pB_1+1$ 或 $q\mid 2pB_2+1$，则 q 为 $pb+1$ 之形状，且

$$m^p \equiv 1(\mathrm{mod}\, q)$$

为最小解。

引理 2.6.2 p 为奇素数，p^p-1 定有 $2np+1$ 型的素因子，其中 $p\nmid n$。

引理 2.6.3 $p^p\equiv 1\ (\mathrm{mod}\, q)$，$q=2np+1$，$p$，$q$ 均为奇素数（以下均假定 p，q 为奇素数）$p\nmid n$，对任意整数 K，则 $K^p \not\equiv p\ (\mathrm{mod}\, q)$。

引理 2.6.4 $p^p\equiv 1\ (\mathrm{mod}\, q)$，$q=2np+1$，且 $p\nmid n$，若 $m^p\equiv 1\ (\mathrm{mod}\, q)$，则 $m\equiv p^d\ (\mathrm{mod}\, q)$，其中 $d\geqslant 0$。

引理 2.6.5 $p^p\equiv 1\ (\mathrm{mod}\, q)$ 为最小解，$q=2np+1$，$p\nmid n$，那么，一定有 K，$K^{2np}\equiv 1\ (\mathrm{mod}\, q)$ 为最小解。

引理 2.6.6 $p^p \equiv 1\ (\bmod q)$，$q = 2np + 1$，$p \nmid n$，$K^{2np} \equiv 1\ (\bmod q)$为最小解，则

$$K^{np} \equiv -\ (\bmod q)$$

亦为最小解，对于 l，$(l, q) = 1$，则存在 b，使 $l^p \equiv K^b p (\bmod q)$。

引理 2.6.7 $p^p \equiv 1\ (\bmod q)$，$q = 2np + 1$，$p \nmid n$，$K^{2np} \equiv (\bmod q)$ 为最小解，则存在 d_1，d_2，使

$K^n \equiv - p^{d_1} (\bmod q)$，$K^{2n} \equiv p^{d_2} (\bmod q)$，且对任意 l，$(l, q) = 1$，当 $n \mid b$ 时，若存在一个 m，使 $l^b \equiv \pm p^m (\bmod q)$，则一定存在一个 m_1，使 $l \equiv \pm p^m (\bmod q)$。

引理 2.6.8 $p^p \equiv 1\ (\bmod q)$，$q = 2np + 1$，$p \nmid n$，$K^{2np} \equiv 1\ (\bmod q)$ 为最小解，当 $n > 1$ 时，若 $n \nmid b$，则存在 c，d，使 $K^{bp} \equiv p^d K^c\ (\bmod q)$，且 $p \nmid dc$，$n \nmid c$。

引理 2.6.9 $p^p \equiv 1\ (\bmod q)$，$q = 2np + 1$，$p \nmid n$，$K^{2np} \equiv 1\ (\bmod q)$ 为最小解，那么

$$1 + K^{bp} + K^{cp} \not\equiv (\bmod q),$$

其中 $b \geqslant 0$，$c \geqslant 0$。

由上面的引理可得如下定理。

定理 2.6.4 $p^p \equiv 1\ (\bmod q)$，$q = 2np + 1$，$p \nmid n$，$K^{2np} \equiv 1\ (\bmod q)$ 为最小解，当 $p \nmid xyz$ 时，若 $q \mid xyz$，则方程

$$x^p + y^p = z^p$$

无正整数解。

定理 2.6.5 $p^p \equiv 1\ (\bmod q)$，$q = 2np + 1$，$p \nmid n$，$k^{2np} \equiv 1\ (\bmod q)$ 为最小解，当 $p \nmid xyz$ 时，若 $q \nmid yz$，则方程

$$x^p + y^p = z^p$$

无正整数解。

然而，上述证明尚未得到数学家的承认。是否存在证明费尔马大定理的初等方法呢？从三百多年来，人们对费尔马大定理的探讨过程来看，用初等数论的方法彻底解决费尔马大定理的可能性十分微小。这里需要说明的是，有的业余数学爱好者，对于欧

拉当年证明 $n=3$ 的情形都没有看过，更不用说对于库姆尔正则素数的有关结果了，他们就宣布已经证明了费尔马大定理。这样所谓的研究成果，只能是白费笔墨。正确的做法应该是在了解和掌握前人研究成果的基础上，去挖掘和发现证明费尔马大定理的新思想、新方法。费尔马大定理虽然问题的叙述十分简单，但要解决它，却十分艰难。没有雄厚而牢固的数学基础，没有丰富的科学想象力，没有献身精神，想靠偶然的机会来解决费尔马大定理，这种想法是天真的，不妥当的。

三、触类旁通

——费尔马大定理与莫德尔猜想

1983 年，西德 29 岁的青年数学家法尔廷斯（Gerd Faltings）证明了数论中长达六十多年之久的莫德尔猜想，在证明的过程中，还解决了泰特猜想和沙发列维奇猜想，从而为解决费尔马大定理迈出了重要的一步。

1993 年，英国数学家安德鲁·维尔斯宣布他已证明了费尔马大定理，1994 年，他给出这个定理严格的证明。

（一）莫德尔猜想

1922 年，英国的数学家莫德尔（Louis J. Mordell）用直接方法研究椭圆曲线（即亏格为 1 的非奇异曲线）上的有理点。之后，他证明了一个椭圆曲线上的有理点形成有限生成交换群，而其仿射曲线只有有限多整点。接着，他提出如下猜想：

设 $f(x,y)$ 是任意一个不可约、有理系数的二元多项式，Q 为有理数域，当 f 的亏格不小于 2 时，最多存在有限对数偶 $x_i, y_i \in Q$，使得 $f(x_i, y_i) = 0$。

这一猜想，就叫做莫德尔猜想。后来人们把这个猜想扩充到了定义在任意数域上面的多项式，并且随着抽象代数几何的出现

又重新用代数曲线来叙述这个猜想，即：

任意定义在数域 k 上，亏格不小于 2 的代数曲线最多只有有限个 k-点。

我们来具体地分析一下这一猜想。

所谓代数曲线，粗略地说，就是在包含 k 的任意域中，$f(x,y)=0$ 的全部解的集合。详言之，域 k 上的 n 维射影空间是 $n+1$ 维向量空间中，通过原点的直线的集合。在选定向量空间的一组基底之后，可以用一直线上的一个非零点的坐标 $(x_0, x_1, \cdots, x_n)$ 来确定这条曲线，即定出射影空间中的一个点。如果 F 是 x_0，x_1，$\cdots$，x_n 的一个齐次多项式，那么 $F(X,Y,Z)$ 为 n 次齐次多项式，其中 n 为 $f(x,y)$ 的次数，且使

$$F(X,Y,1)=f(x,y);$$

这样的 F 是唯一的。这时，F 在二维射影空间中的零点集合包含了 $f(x,y)=0$ 的解，但是还有其他的对应于 $Z=0$ 平面上直线上的点，使得 $F=0$。

f 的亏格的公式为

$$\text{亏格}=(n-1)(n-2)/2-\sum_p v_p,$$

其中 Σ 是对于满足

$$\frac{\partial F}{\partial X}(p)=\frac{\partial F}{\partial Y}(p)=\frac{\partial F}{\partial Z}(p)=0$$

的射影点 p 取的，且 $v_p>1$。这样些点叫做零点轨迹的奇点。如费尔马曲线 x^n+y^n-1 没有奇点，因此其亏格为 $(n-1)(n-2)/2$。

从亏格不小于 2 的要求，可推出其次数应不小于 3。为什么次数不小于 3 呢？下面予以考察。当 $n=1$ 时，$f=ax+by+c$ 显然有无穷多个解。当 $n=2$ 时，f 可能没有解，例如，在有理数域上，$f=x^2+y^2+1$ 就没有解；但是，如果 f 有一个解就必定有无穷多个解。我们可用几何方法来证明这一点。记 $B=\{f$ 的解集合$\}$，设点 $p\in B$，令 L 为不过点 p 的一直线。点 $Q\in K$，

且在 L 上，直线 PQ 与 B 交于另一点 R，当 Q 在 L 上取遍无穷多个 K-点时，R 点的集合就是 f 的 K-解的无穷集合。例如，把这种方法用于 $f=x^2+y^2-1$，给出参数公式。

$$x=\frac{t^2-1}{t^2+1}, y=\frac{2t}{t^2+1}。$$

当 F 为三次非奇异曲线时，其解集合是一个解，即一条所谓椭圆曲线。这个群的规则为：经过 P，Q 点的直线交零点集于第三个点 R，这个 R 就是 $-P-Q$。选定一个适当的原点 P_0，点 R 的 2 倍。假定这个起始点 R 对群来说，不是有限阶，则上述的迭代产生了一个解的无穷集合。

由上可知，为什么猜想中要求其亏格大于或等于 2。

（二）解不定方程的一般性问题

诸项具有相同次数的不定方程叫做是齐次不定方程；否则，叫做非齐次不定方程。对于齐次不定方程，如果向量 x 是解，则对于每一个常数 c，cx 也是解。因此，这时将彼此相差一个非零常数因子的两个解看成是等同的，对于齐次方程我们只求不等同的解。由一个非齐次 n 次方程总可以变成一个与它相关联的齐次方程，方法是：先将 n 个变量集合，再增加一个变量，使之变成 $n+1$ 个变量，然后，把原方程的每一项乘以新变量适当的方幂，使每一项的次数均为 n 次。例如，对于非奇次方程 $y^2=x^2\ (x+1)$ 变成齐次方程 $y^2z=x^2\ (x+z)$。于是原来非齐次方程的解，对应于齐次方程新变量等于 1 的解。在实际应用中，我们对于非齐次方程情形求整数解，而对齐次方程情形求有理数解。n 个变量的非齐次方程和 $n+1$ 个变量的齐次方程的关系就像仿射几何与射影几何的关系一样。

利用几何，是研究不定方程的另一条途径，每个不定方程的复数解的轨迹定义出一个复的代数簇。如果不定方程是非齐次

的，其全部解形成一个仿射代数簇；若是齐次的，其全部解形成一个射影代数簇。由此可知，非齐次方程和相应的齐次方程的关系恰好对应于仿射代数簇和它的适当的射影空间取射影闭包（即添加一个无穷远点）所得的射影代数簇。

于是，研究不定方程的解变成了 求由系数属于某数域的方程组定义的复代数簇上的整数点或者有理点。例如费尔马所研究的方程是射影平面中亏格为 $\frac{1}{2}(n-1)(n-2)$ 的非奇异曲线的方程。人们希望几何学能使某些不定方程问题更容易处理。第一步显然是把注意力局限于对应于曲线的那些不定方程。利用这种方法，可以把不定方程的求解问题重新叙述成寻求曲线上有理点和整点的问题。

另一个相关问题是定义在有限域上的光滑射影代数簇，这样一个代数簇在各个不同有限扩域中的有理点数可以表示成一个母函数，叫做此代数簇的 zeta 函数。关于这个 zeta 函数的性质魏尔（Weil）有一系列猜想，这些猜想使我们在理论上可以计算这个 zeta 函数。

（三）几个重要结果

1. 曲线的沙发列维奇（Shafarevich）猜想

1962 年，沙发列维奇提出如下猜想：

设 K 是 Q 的有限扩张，S 是 K 的一个有限的位集合，则亏格 $g\geqslant2$，并且在 S 外具有好的约化的光滑曲线 X/K，只有有限多个同构类。

1968 年，泊辛（A.N.Parshin）证明了莫德尔猜想可由曲线的沙发列维奇猜想推出来。比如，假设齐次多项式 F 为非奇异的，且 $K=Q$。消去 F 的系数的分母之后，可设 F 的系数为无

因子的整数；那么，我们可以对某个素数 p，考虑其同余方程。如果其偏导 mod$p$ 没有公因子，则方程 $F(X, Y, Z) \equiv 0 (\bmod p)$ 给出域 Z/pZ 上的非奇异曲线，这时，就称原来的曲线在 P 具有良约化。例如，费尔马曲线 $X^3+Y^3+Z^3$ 在素数 3 没有良约化。曲线 $Y^2Z=X^3-17Z^3$ 在 Q 上为非奇异，而 mod2，3 或 17 时，它都是奇异的。

从猜想之间的关系考虑问题，是解决猜想的一个途径。由 Parshin 的研究结果表明，要证明莫德尔猜想就归结到只要证明曲线的沙发列维奇猜想即可，这一猜想正是法尔廷斯证明的。

2. 阿贝尔簇的沙发列维奇猜想

用“过渡到 Jacobi 簇”方法去处理曲线的问题常常是方便的。从数论的观点看这就是由椭圆曲线到阿贝尔簇的推广。由沙发列维奇提出、由 Faltings 证明的，实则是对于阿贝尔簇的更强形式的沙发列维奇猜想。

Jacobi 簇的构造可用黎曼曲面的知识来阐述，也可以代数的方式进行，这就产生了定义在域 K 上的一个阿贝尔簇，其中 K 即原曲线的定义域。阿贝尔簇是一个具有群结构的、射影空间的闭连通子簇。它定义在 K 上即表明它是系数在 K 的一组多项式的零点集。阿贝尔簇的沙发列维奇猜想为：

对于具有给定维数 g，极化次数 $d>0$，并且在位的有限集 S 之外有好的约化的极化阿贝尔簇，其 K-同构类集合是有限的。

由 Torelli 定理知，若阿贝尔簇的沙发列维奇猜想真，则曲线的沙发列维奇猜想亦真。

3. 有界高度原理

费尔马用无穷递降法证明了费尔马大定理 $n=4$ 的情况，现在我们从另一个角度来叙述这一过程，从而引导出新的结果。费

尔马在证明不定方程 $x^4+y^4=z^4$ 没有非零整数解时，对每个 (p,q,r) 给出一个高，例如 $|r|$，其中 (p,q,r) 满足这个方程。无穷递降的方法，使他能够证明，当 (p,q,r) 在曲线上并且有正的高时，则有一组新的 (p',q',r') 也在曲线上，而它有较小的高，即 $0<|r'|<|r|$。从而得到在 $x^4+y^4=z^2$ 上不存在 $(0,0,0)$ 以外的整数点。法尔廷斯将这一方法进行推广，就得到阿贝尔簇的高度理论。

法尔廷斯从半阿贝尔簇理论开始，这是域上阿贝尔簇概念到概型上代数簇的推广，他讨论了关于稳定曲线和主极化阿贝尔簇的参模空间的一些结果。然后，他发展了可利用于半阿贝尔簇的高度理论。从而得到了有界高度原理。这个原理是说：

对于给定的常数 c，高度小于或等于 c 并具有某些性质的半阿贝尔簇的同构类数是有限的。

这个原理是法尔廷斯证明莫德尔猜想的基础。科学研究既要继承更要开拓，要在已有成果的基础上，开发出新的东西出来，阿贝尔簇的高度理论就是从无穷递降法通过科学家的科学想象力，经过细心的研究而扩展出来的。

4. 同源下高的行为

由定义知，阿贝尔簇的一个同源是代数群之间的一个具有有限核的满同态。

在许多情形下，同源是阿贝尔簇的等价性的“正确”概念。因而许多涉及阿贝尔簇的问题用簇与它的同源像的等价性去处理很有好处。为此，必须研究一下高在同源下是怎样变化的。

一个同源在 $\bmod p$ 约化时可能出现的分歧，我们以下例说明之。考虑由 $X^3+Y^3-1=0$ 定义的一维阿贝尔簇。如果取 $(1,0)$ 为群结构的原点，于是乘以 -2 可几何地描述为过 p 的曲线的切线与曲线交出的第三个点，曲线在点 (x,y) 的切线由 $x^2X+y^2Y-1=0$ 定义，所以乘以 -2，可由坐标变换

$$(x,y)\longrightarrow\left(\frac{x^4-2x}{1-2x^3},\frac{x^4-2y}{1-2y^3}\right)$$

推出。这个映射的 mod 2 约化在点 (1,0) 上有垂直切线。同源的映射度的对数的一半再减去某个修正项给出了在同源下高的改变值，这个修正项度量出这一类分枝。

5. 泰特猜想

定义在域 K 上的阿贝尔簇 A 上，其 Tate 模记为 $T_l(A)$。如果 B 是另外一个阿贝尔簇，则可以去比较 K 上阿贝尔簇间的映射群 Hom (A, B) 和相应的伽罗华群 G-映射群 $\mathrm{Hom}_G(T_L(A), T_L(B))$，其中后者表示与 G 作用可交换的同态构成群。于 1967 年 J. Tate 提出如下猜想：

$T_L(A)$ 上的伽罗华表示为半单纯的，且当 K 在其素域上为有限生成时，有

$$\mathrm{Hom}(A,B)\otimes Z_i\approx\mathrm{Hom}_G(T_L(A),T_L(B))$$

1974 年，Zarhin 的 Tate 方法与论断导致 Tate 猜想的证明。

（四）莫德尔猜想的证明

现在把法尔廷斯对莫德尔猜想的证明以方块箭头图的形式指示如下，其中六边形中表示法尔廷斯的主要研究结果，长方块中表示已有的结果。

法尔廷斯之所以证明了莫德尔猜想，除了他的科学的想象力之外，还与他全面研究与了解前人对这个问题所进行的工作是分不开的。早在三四十年代，随着阿贝尔簇理论的发展，人们就逐渐认识到莫德尔猜想应当理解成为关于包含亏格不小于 2 的曲线的某种阿贝尔簇的一个命题。到了 60 年代 S. Lang 把莫德尔猜想给出一种重述形式，即阿贝尔簇上的一条曲线与此阿贝尔簇的每个有限生成子群只有有限多个公共点。S. Lang 还证明了，莫

德尔猜想也可视作是关于曲线代数簇的一个猜想。60年代后期，Deligne等人或者通过分析特殊曲线簇，或者运用高度理论来探索莫德尔猜想的证明。60年代一个重大突破是证明了莫德尔猜想在函数域上的等价命题。Neron证明了，由两个莫德尔猜想能够建立起 Q 的任何有限生成扩域上的相应猜想。Manin和Grauert都致力于证明莫德尔猜想在函数域上的类似命题。1963年沙发列维奇和1968年Parshin集中研究沙发列维奇猜想和它与莫德尔猜想在函数域上类比的联系。法尔廷斯在这些前人工作的启发下，证明了莫德尔猜想。

通过莫德尔猜想的证明，使我们看到一个猜想的解决，还受一定的科学发展水平的制约。一个不定方程的难题——莫德尔猜想，用发展起来的代数几何方法所解决，如果仅限于用初等数论的方法，肯定说现在还不知道从哪个方向去攻克这一难题。法尔廷斯证明莫德尔猜想的意义还在于我们看到它与解决阿贝尔簇理论中许多重要猜想之间的联系。这也就启发我们，解决一个数学难题不能孤立去考虑，还应该考虑难题之间的联系，一个难题解决了，另一个难题也就可能随之而解决了。

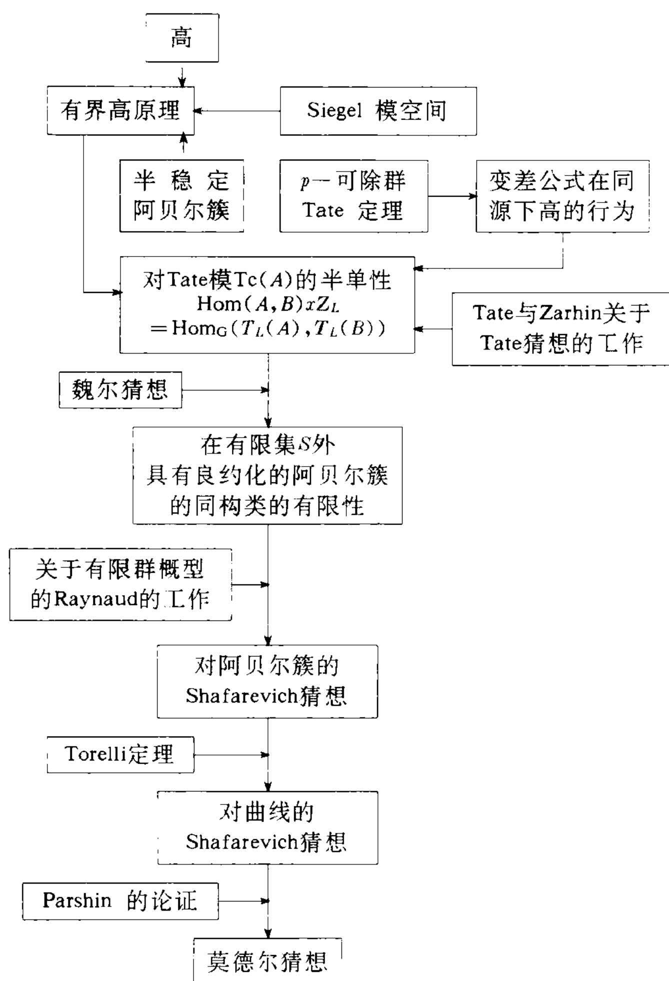
（五）从莫德尔猜想到费尔马大定理

法尔廷斯已经证明了莫德尔猜想，这就是说，如下结论是对的：

设 $F(x,y)$ 是两个变量 x,y 的有理系数多项式，那么当曲线 $F(x,y)=0$ 的亏格不小于2时，方程 $F(x,y)=0$ 至多有有限组有理解。

特别，对于一个特殊的多项式 x^n+y^n-1 当 $n\geqslant 4$ 时，其亏格 $(n-1)(n-2)/2>2$，故知方程 $x^n+y^n-1=0(n\geqslant 4)$ 至多有有限组有理解，从而推知它所对应的齐次方程 $x^n+y^n=z^n(n\geqslant 4)$ 只有有限多个整数解。如果我们再能进一步证明这有限多

莫德尔猜想证明示意图



个整数解为空集，那么费尔马大定理就得证了。通过费尔马大定理与莫德尔猜想之间的联系，使我们再一次看到，考虑猜想之间的联系是解决猜想的一个重要思维途径。

（六）模曲线和费尔马大定理

椭圆曲线是最典型的三次曲线。一般的三次平面曲线方程为

$$ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + ky + l = 0.$$

经过坐标变换，再进一步简化，上述方程就可以转化成椭圆曲线的标准形式：

$$y^2 = x^3 + a_1x + a_2。$$

德国数学家符莱把椭圆曲线与费尔马大定理联系起来。1985年，他证明了，如果费尔马方程

$$x^p + y^p = z^p (p \text{ 为不小于 5 的素数})$$

有非零解(a, b, c)，即

$$a^p + b^p = c^p,$$

则可设计一条椭圆曲线

$$y^2 = x(x + b^p)(x + c^p),$$

其中不妨假定 a，b，c 为互素的非零整数，显然它是有理数域上的椭圆曲线。又由于方程右端没有重根，这种曲线就称为符莱曲线。对于其中的 b，c 作些限制，就得到半稳定的椭圆曲线。

椭圆曲线可以用一些函数进行参数表示。如果参数表示所用的函数能用模形式，就称之为模椭圆曲线，简称模曲线。任一椭圆曲线都是模曲线，这就是谷山一志村猜想。这个关系于1986年是李贝由塞尔的猜想证明的。这样一来，要证明费尔马大定理，只需证明对半稳定椭圆曲线，谷山一志村猜想成立即可。英国数学家安德鲁·维尔斯知道这一消息之后，经过七年的苦心研究，终于在1993年6月23日通过塞尔姆群构造欧拉系而取得了突破。1994年，他以海克代数的结构理论为工具，严格证明了费尔马大定理，他将定理证明的过程整理成论文《模曲线和费尔马大定理》于10月14日送交当代最权威的数学杂志——普林斯

顿的《数学年刊》，该刊于1995年5月刊发了这篇论文。

（七）费尔马大定理获证之后

费尔马大定理经过357年的努力，终于被攻克了，这标志着一个时代的结束。由于数学家使用了强有力的代数几何方法，使人们对零散的、各种各样的不定方程问题有了一个系统的了解。通过曲线的不变量——亏格来看问题，对于二个未知数的非齐次方程和三个未知数的齐次方程已经有了基本的了解，它们的整数解与有理数解原则上是清楚的。

从不定方程来看，这仅是沧海之一粟，大部分问题还有待于进一步解决。费尔马大定理可以看成是求解不定方程

$$x^n + y^n + z^n = 0$$

的整数解，一个最自然的推广是求不定方程

$$x^n + y^n + z^n = m$$

的整数解。这是非齐次形式的推广，还可以从增加未知数数目，向齐次形式推广。此外，也可以从增加系数，增加方程的数目方向予以推广。由此可以看出，许许多多数论猜想就是这样由简单到复杂无休无止地产生出来，一个猜想解决了，更多的猜想又在等着你，数学家总有解决不完的猜想。

四、一步之遥

——哥德巴赫猜想

哥德巴赫猜想是解析数论的一个中心课题。这一猜想从提出到现在已经二百多年了，但至今没有被证明。为了解决这一问题，许多数学家付出了艰苦的劳动，并取得了一系列成果。我国著名数学家陈景润解决了哥德巴赫猜想 1+2 的问题，被数学家誉为“陈氏定理”。到目前为止，这是对哥德巴赫猜想研究的最好结果。如果解决了哥德巴赫猜想 1+1 的问题，那么哥德巴赫猜想就彻底解决了。目前解决这一猜想，虽然只有一步之差，但这一步究竟如何迈出，又何时达到终点，这是数学家当前难以预料的问题。

（一）猜想的提出

在两个数相加中，我们会遇到如下的关系：

$3+7=10$，

$3+17=20$，

$13+17=30$，

$17+23=40$，

$13+37=50$。

我们来分析一下上述等式有什么相似之处。很自然地会发现：等

式右边的数都是偶数，等式左边的两个数都是奇素数。我们已经知道两个奇素数之和必定是一个偶数。反过来，我们要问：任一个偶数都可以分拆为两个奇素数之和吗？我们再作一些观察。第一个等于两个奇素数之和的偶数为

$$6=3+3。$$

接下去为

$$8=3+5,$$
$$10=3+7=5+5,$$
$$12=5+7,$$
$$14=3+11=7+7,$$
$$16=3+13=5+11,$$
$$18=5+13=7+11,$$
$$20=3+17=7+13,$$
$$22=3+19=5+17=11+11,$$
$$24=5+19=7+17=11+13,$$
$$26=3+23=7+19=13+13,$$
$$28=5+23=11+17,$$
$$30=7+23=11+19=13+17。$$

通过上述个例观察，可知这些偶数都可分拆成两个奇素数之和，于是，由特殊到一般，我们可提出如下猜想：

(A) 任何$\geqslant 6$的偶数都是两个奇素数之和。

对于偶数可提出上述判断，对于奇数是否也可提出类似结论呢？显然，奇数不能分拆成两个奇素数之和。既然两个不行，那么分拆成三个奇素数之和，行吗？通过下面的实例进行观察：

$$9=3+3+3,$$
$$11=3+3+5,$$
$$31=3+5+23=3+11+17=5+7+19=5+13+13,$$

由特殊到一般，于是可猜想：

(B) 任何$\geqslant 9$的奇数都是三个奇素数之和。

上述规律是否有普遍性？德国著名数学家哥德巴赫（Christian Goldbach，1690—1764）对这个问题发生了浓厚的兴趣，但是，他不敢肯定其正确性。于是，他于1742年写信给当时的数学权威欧拉，就此问题进行请教。他问欧拉：是不是每个偶数都是两个素数之和，每个奇数都是三个素数之和？欧拉于1742年6月30日回信说：他验算到100多，发现是对的，但不能给出一般性的证明。

到1770年，华林首次把这个问题以猜想的形式写在书中，并公诸于世。由于当时把1也看成素数，所以问题提的不太确切。确切的提法是上面所述的猜想（A），（B）。猜想（A）叫做偶数哥德巴赫猜想，猜想（B）叫做奇数哥德巴赫猜想。易知，由（A）成立，可推出（B）成立。事实上，如果（A）成立，设 N 是一个大于7的奇数，那么 $N-3$ 就是一个 $\geqslant 6$ 的偶数，据（A），有

$$N - 3 = p_1 + p_2,$$

其中 p_1，p_2 为奇素数。因此，

$$N = p_1 + p_2 + 3$$

是三个奇素数之和。从而命题（B）成立。这样一来，只要解决（A），（B）也就随之而解决了。

（二）悲观的预言与惊人的成果

从18世纪40年代哥德巴赫猜想的提出，到19世纪末，许多数学家都对这一猜想进行了研究，但在这一百五十多年中，并没有得到任何实质性的结果和提出有效的研究方法。只是对一些数值作了进一步的验证，使猜想变得更加可信，增加了它的合理性。另外还提出一些简单的关系式和一些新的推测。在这一期间数学家虽然对哥德巴赫猜想的探讨作了极大的努力，但是由于用来解决这一问题的数学理论还没有发展到这个地步，因此进展缓

慢。与此同时，由于欧拉、高斯、狄里克雷、黎曼、哈达马等著名数学家的工作，使数论和函数论得到了空前的丰富和发展，特别是分析与数论相结合，在数论中引入了分析的方法，这就为20世纪对这一猜想的研究提供了强有力的工具。在这一百五十多年中，研究哥德巴赫猜想没有什么进展，这从反面说明，解决数学难题，得有足够的数学基础知识。如果连初等数论尚没有弄明白，更不用说解析数论和函数论，就想一下子证明哥德巴赫猜想，肯定说是异想天开，白费力气。

1900年，在巴黎召开的第二届国际数学会上，德国著名数学家希尔伯特提出了数学中著名的23个问题，哥德巴赫猜想就是第八个问题的一部分。在这之后的十多年，对哥德巴赫猜想的研究并未取得进展。1912年，德国数学家朗道（E. Landau, 1877—1938）在英国剑桥召开的第五届国际数学会上悲观地说：即使要证明下面较弱的命题（C），也是当代数学家所力不能及的：

（C）存在一个正整数 k，使每一个 $\geqslant 2$ 的整数都是不超过 k 个素数之和。

1921年，英国数学家哈代（G.H.Hardy, 1877—1947）在一次数学会上也谈到：哥德巴赫猜想，可能是没有解决的数学问题中的最困难的一个。

在解决这一难题的过程中，数学家看到其艰巨性，特别在亲自尝试过程中，其体会更为深刻。但勇于探索的人们，并没有望而止步，而是不断地为之拼搏，努力地从前人研究所走过的道路上，去挖掘解决哥德巴赫猜想可能取得成果的潜在思想。正当一些数学家对此猜想感到无能为力时，数学家却开始从不同的方向上取得了一系列惊人的成果。这些成果的取得，不仅为解决哥德巴赫猜想开拓了途径，而且还有力地推进了数论和其他数学学科的发展。

（三）圆　　法

19 世纪中叶，狄利克雷和黎曼把分析方法移植到数论中来，从而使数论得到了空前的发展，使一些一筹莫展的问题，有了解决的希望。从 1920 年开始，英国数学家哈代和李特渥德（Littlewood，1885—1977）系统地开创与发展了堆垒素数论中的一个崭新方法。1923 年发表论文专论哥德巴赫猜想。这一新方法的思想孕育在 1918 年哈代和印度数学家拉曼牛建（Ramanujan）的文章中，后来人们就称这个新方法为 Hardy-Littlewood-Ramanujan 圆法。这个方法，对于哥德巴赫猜想来说，就是把数论中离散的问题归结到连续问题来处理。其基本思想是：设 m 为整数，由于积分

$$\int_0^1 e(ma)da = \begin{cases} 1, & m = 0; \\ 0, & m \neq 0; \end{cases}$$

其中 $e(x) = e^{2\pi i x}$，所以方程

$$N = p_1 + p_2, p_1, p_2 \geqslant 3 \tag{1}$$

的解数为

$$D(N) = \int_0^1 S^2(a, N)e(-Na)da; \tag{2}$$

方程

$$N = p_1 + p_2 + p_3, \quad p_1, p_2, p_3 \geqslant 3 \tag{3}$$

的解数为

$$T(N) = \int_0^1 S^3(a, N)e(-Na)da, \tag{4}$$

其中

$$S(a, N) = \sum_{2 < p \leqslant N} e(ap)。 \tag{5}$$

这样一来，猜想（A）就归结为要证明：对于偶数 $N \geqslant 6$，则有

$$D(N) > 0;$$

猜想（B）就归结为要证明：对于奇数 $N \geqslant 9$，则有

$$T(N) > 0。$$

于是，哥德巴赫猜想就转化为讨论关系式（2），（4）中的积分了。因而这就需要研究由（5）所确定的以素数为变数的三角和。（5）的性质知道了，其积分的值也就求出来了。（5）有什么性质呢？我们猜测：当 a 和分母“较小”的既弱分数“较近”时，$S(a,N)$ 就取“较大”的值；而当 a 和分母“较大”的既约分数“接近”时，$S(a,N)$ 就取“较小”的值。这样我们就可把积分区间分成两部分，在其中的一部分，是积分的主要项，积分易求出来，而另一部分，是积分的次要项，积分值可忽略不计。这就是圆法的主要思想。下面就此稍加具体的说明。

设 M，τ 为两个正数，

$$1 \leqslant M \leqslant \tau \leqslant N。$$

考虑法列数列

$$\frac{a}{q},(a,q)=1,0 \leqslant a < q, q \leqslant M。$$

并设 $E(q,a)=\left[\frac{a}{q}-\frac{1}{\tau},\frac{a}{q}+\frac{1}{\tau}\right]$ 以及

$$E_1=\bigcup_{1 \leqslant q \leqslant M} \bigcup_{0 \leqslant a \leqslant q} E(q,a),$$

$$E_2=\left[-\frac{1}{\tau},1-\frac{1}{\tau}\right] \Big| E_1。$$

易证，当

$$2M^2 < \tau$$

时，所有的小区间 $E(q,a)$ 是两两不相交的。称 E_1 为基本区间，E_2 为余区间。如果一个既约分数的分母不超过 M，我们就说它的分母是“较小”的；否则，就说是“较大”的。如果两个点之间的距离不超过 τ^{-1}，我们就说是“较近”的。显然，当 $a \in E_1$ 时，它就和一分母“较小”的既约分数“接近”。当 $a \in E_2$ 时，可以证明，它一定和一分母“较大”的既约分数“接近”。

这样利用法列数列就把积分区间 $\left[-\frac{1}{\tau}, 1-\frac{1}{\tau}\right]$分成了圆法所要求的两部分 E_1 和 E_2。

为方便起见，我们把积分区间 $[0,1]$ 改为$\left[-\frac{1}{\tau}, 1-\frac{1}{\tau}\right]$。这样一来，(2)，(4) 的积分就分成两部分，即

$$D(N) = \int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}} S^2(a,N)e(-Na)da = D_1(N) + D_2(N), \tag{6}$$

其中 $D_i(N) = \int_{E_i} S^2(a,N)e(-Na)da,\ i = 1,2$

$$T(N) = \int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}} S^3(a,N)e(-Na)da = T_1(N) + T_2(N), \tag{7}$$

其中 $T_i(N) = \int_{E_i} S^3(a,N)e(-Na)da,\quad i = 1,2$。

圆法就是要计算出 $D_1(N)$ 及 $T_1(N)$，并证明其为 $D(N)$，$T(N)$ 的主要项，而 $D_2(N)$，$T_2(N)$ 分别作为其次要项。

如果不加任何条件限制，难以计算出 $D(N)$，$T(N)$ 的渐近式。这样一来，就想到把考虑问题的范围缩小，于是 1923 年，哈代、李特渥德取得了第一个突破，他们证明了如下结论。

在弱型广义黎曼猜想成立的前提下，每个大奇数一定可表为三个奇素数之和，且有渐近公式

$$T(N) = \frac{1}{2}R_3(N)\frac{N^2}{\log^3 N}, N \to \infty, \tag{8}$$

其中 $R_3(N) = \prod_{p\mid N}\left(1-\frac{1}{(p-1)^2}\right)\prod_{p\nmid N}\left(1+\frac{1}{(p-1)^3}\right)$。 (9)

对于偶数又怎样呢？他们猜测有

$$D(N) = R_2(N)\frac{N}{\log^2 N}, N \to \infty,$$

其中 $$R_2(N) = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right)\prod_{\substack{p\mid N\\ p>2}}\frac{p-1}{p-2}。$$

对于一个大的猜想，在一个较长的时间解决不了，我们可以将猜想进行转化，可以对猜想加上前提条件，先得到一个带有假设性的结果，或者加上前提条件，再提出新的猜测，然后对这新的猜测进行探求。

虽然，哈代、李特渥德没有证明任何无条件的结果，但是他们在证明有条件的结果所创造的圆法，为人们指明了一个有成功希望的研究方向。正如他们自己所说："我们借助于堆垒数论中新的超越方法来攻这个问题，我们没有解决它。甚至我们也没有证明任何数是 1000000 个素数之和。……然而，我们证明了这个问题不是攻不动的，……。"这就是说，他们创造的圆法，消除了人们对研讨哥德巴赫猜想的悲观情绪，增进了解决此问题的必胜信心。事实上，圆法为人们解决哥德巴赫猜想找到了一个有效途径，为下一个突破创造了良好的条件，同时，它在解决数论中的其他难题中也发挥了积极作用。

1937 年，苏联数学家维诺格拉朵夫（И. М. Виноградов）在"圆法"的基础上，再加上他独创的"三角和估计方法"，去掉了弱型广义黎曼猜想的前提，证明了：每一个充分大的奇数都是三个奇素数之和，且有渐近公式 (8) 成立。后来，有人用别的分析方法，也"无前提"地证明了这个结果。这些大奇数究竟有多大？它比 1 后面带上几十万个零还要大。这虽然是一个天文数字，但剩下的数总是有限的，原则上总是可以一一验证的。由无限转化到有限，这是一个重大突破，因此猜想 (B) 算是基本解决了。这一结果通常叫做哥德巴赫—维诺格拉朵夫定理，简称三素数定理。

维诺格拉朵夫是怎样证明三素数定理的呢？

1935 年，Page 证明了

定理 4.3.1 设整数 $q \geqslant 3$，则对所有实的非主特征 $X \bmod q$[①]，当 $\sigma \geqslant 1-\frac{C_4}{\sqrt{q}\log^4 q}$ 时，有

$$L(\sigma, x) \neq 0$$[②]。

定理 4.3.2 设整数 $q \geqslant 1$，x 是模 q 的实特征，则对任一给的 $\varepsilon > 0$，一定存在一个常数 $c = c(\varepsilon) > 0$，使得 $L(s, x)$ 的非实零点 β，满足

$$\beta \leqslant 1 - \frac{c(\varepsilon)}{q^{\varepsilon}}。$$

由上述两个定理可推出相应的算术级数中素数分布的如下两个定理。

定理 4.3.3 设 $x \geqslant 2$，则对于任意固定的正数 $A > 1$，及任意的整数 q，l，

$$1 \leqslant q \leqslant \log^A x, (l, q) = 1,$$

有渐近公式

$$\psi(x; q, l) = \frac{x}{\phi(q)} + o(xe^{-c_2\sqrt{\log x}}),$$

$$\pi(x; q, l) = \frac{Lix}{\phi(q)} + o(xe^{-c_2\sqrt{\log x}})$$

成立，其中常数 c_2 依赖于 A，且 o 常数是一绝对常数，c_2 是不能实际计算出的常数。

定理 4.3.4 设 $x \geqslant y > 3$，则对所有的模 $q \leqslant y$，可能除去一些“例外模” q——这些 q 一定是某一个可能存在的 q_0（$q_0 \gg \log^2 y(\log\log y)^{-s}$）的倍数 —— 以外，当 $(q, l) = 1$ 时，有如下式成立：

$$\pi(x; q, l) = \frac{Lix}{\phi(q)} + o(xe^{-c_3\sqrt{\log x}}) + o(xe^{-c_3\frac{\log x}{\log y}}),$$

---

① 特征 $x(h)$ 是属于模 q，记作 $x(h) \bmod q$，模 q 的特征 $x(h)$ 称模 q 的主特征，其他的所有特征都称为非主特征。

② $L(\sigma, x)$ 为 L 函数。

其中，大 o 常数及 c_3 都是绝对的可计算的常数。

上述两个定理之一可推出如下结果。

定理 4.3.5 对于奇数 N 表为三个奇数之和的表法个数 $T(N)$ 有渐近公式

$$T(N)=\frac{1}{2}R_3(N)\frac{N^2}{\log^3 N}+o\left(\left(\frac{N^2}{\log^4 N}\right),\right.$$

其中，$R_3(N)$ 为 (9) 式所示，且 $R_3(N)>\frac{1}{2}$。

维诺格拉朵夫成功地创造了素变数三角和估计方法，证明了哈代、李特渥德关于三角和 $S(a,N)$ 性质的猜测，即，他证明了：适当选取 M，τ，当 $a\in E_2$ 时，有

$$S(a,N)\ll\frac{N}{\log^3 N}。\tag{10}$$

由此易推出

$$T_2(N)\ll\frac{N}{\log^3 N}\int_0^1|S^2(a,N)|da\ll\frac{N^2}{\log^4 N}。$$

这就表明 $T_2(N)$ 对 $T_1(N)$ 来说是可以忽略的次要项，从而就证明了三素数定理。

维诺格拉朵夫处理基本区间 E_1 上的积分用的是分析方法，而处理余区间 E_2 上的积分用的是非分析方法。这种方法上的不一致性就导致了数学家去探索用分析方法得到线性素变数三角和 $S(a,N)$ 的估计式(10)。1945 年，Ю. B. 列尼克提出了所谓 L 函数零点密度估计方法，他利用这个方法证明估计式（10），从而使三素数定理给出一个完全有意义的分析法证明。他的这一方法解决了解析数论中的许多问题。这种协调一致的方法上的思考是一种数学美的追求。由于数学美的驱使，促使数学家去创造新的方法。而创造的新方法，又为解决更广泛的一类问题提供新的工具。由此看来，追求数学上的美是丰富和发展数学的一种不可忽视的动力。

维诺格拉朵夫创造的估计三角和的方法是解析数论中的强有

力的工具，应用这一方法，获得了解析数论中许多重要结果，为数论的发展起到了重要的推进作用。

三素数定理证明了。接下去一个很自然的想法就是再推广这一结果。1938 年，我国著名的数学家华罗庚证明了如下定理。

定理 4.3.6 对任意给定的整数 k，使每一个充分大的奇数都可表为

$$p_1+p_2+p_3^k,$$

其中 p_1，p_2，p_3 为奇素数。

特别 当 $k=1$ 时，就是三素数定理。

另一个自然的想法就是在定理中再加些限制条件进行讨论。在前面的例子中，我们已经看到，一个奇数分成三个奇素数之和是不唯一的，同一个奇数，有的分解成的三个奇素数相差比较大，有的分解成的三个奇素数差不多一般大。因此，可提出如下问题：

一个充分大的奇数表为三个几乎相等的奇素数之和。

本世纪 50 年代开始研究这一问题，答案是肯定的。

对于一个猜想，如果加上限制条件，还难以推出结论，就把结论再减弱，这也是解决猜想的一种重要途径。1923 年，哈代、李特渥德，得到了如下的假设性结果：如果广义黎曼猜测成立，那么几乎所有的偶数都能表为二个奇素数之和，即有如下定理。

定理 4.3.7 若以 $E(x)$ 表示不超过 x 且不能表为二个奇素数之和的偶数个数，在 GRH①下，则有

$$E(x)\ll x^{\frac{1}{2}+\varepsilon},$$

其中 ε 为一任意小的正数。

维诺格拉朵夫证明三素数定理之后不久，库尔浦特（J.G.

---

① 所有 $L(s,x)$ 的非显明零点亦都位于直线 $R.S=\frac{1}{2}$ 上，这就是广义黎曼假设，简记作 GRH。

Corput)，楚德可夫(E.C.Чудаков)，艾斯特曼(T.Estermann)，赫尔泊龙(H.Heilbronn)及华罗庚，利用维诺格拉朵夫的思想方法，几乎同时证明了如下定理。

定理 4.3.8 对于任给的正数 A，有

$$E(x) \ll \frac{x}{\log^A x}。$$

下面把这一定理的证明思想简述如下。

我们把能够表为二个奇素数之和的偶数称为哥德巴赫数，而把不能够表为二个奇素数之和的偶数称为非哥德巴赫数。所有不超过 x 的非哥德巴赫数所组成的集合及其个数均用 $E(x)$ 表示。$E(x)$ 亦称作哥德巴赫数的例外集合。于是，对于偶数的哥德巴赫猜想就是要证明：

当 $x \geqslant 4$ 时，有

$$E(x) = 2。$$

设 x 为充分大的正数，以 $D(n,x)$ 表示方程

$$n = p_1 + p_2,\ 2 < p_1 \leqslant x,\ 2 < p_2 \leqslant x$$

的解数。显然，当 $n \leqslant 4$ 时，或 $n > 2x$ 时，恒有

$$D(n,x) = 0。$$

同时，若

$$D(n,x) > 0,$$

则 n 一定是哥德巴赫数。

设 $S(a,x) = \sum\limits_{2 < p \leqslant x} e(ap)$，则显然有

$$D(n,x) = \int_0^1 S^2(a,x)e(-an)da。$$

设 $M = \log^\lambda x$，$\tau = x^{-1}$，$\lambda \geqslant 9$ 为待定正常数，对于这样的 M，τ，可以确定基本区间 E_1 和余区间 E_2。于是，有

$$\begin{aligned} D(n,x) &= \int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}} S^2(a,x)e(-an)da \\ &= D_1(n,x) + D_2(n,x) \end{aligned}$$

其中
$$D_1(n,x)=\int_{E_1}S^2(a,x)e(-an)da$$
$$=\int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}}S_1^2(a,x)e(-an)da,$$
$$S_1(a,x)=\begin{cases}S(a,x), & a\in E_1,\\ 0 & a\in E_2,\end{cases}$$
$$D_2(n,x)=\int_{E_2}S^2(a,x)e(-an)da$$
$$=\int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}}S_2^2(a,x)e(-an)da,$$
$$S_2(a,x)=\begin{cases}S(a,x), & a\in E_2,\\ 0 & a\in E_1。\end{cases}$$
如果能够证明
$$|D_1(n,x)|>|D_2(n,x)|, \tag{11}$$
那么一定有
$$D(n,x)>0,$$
因而 n 就一定是哥德巴赫数。利用维诺格拉朵夫证明三素数定理的思想及如下关系式
$$\sum_n |D_1(n,x)|^2=\int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}}|S_1(a,x)|^4da$$
$$=\int_{E_1}|S_1(a,x)|^4da,$$
$$\sum_n |D_2(n,x)|^2=\int_{-\frac{1}{\tau}}^{1-\frac{1}{\tau}}|S_2(a,x)|^4da$$
$$=\int_{E_2}|S_2(a,x)|^4da,$$
就可证明：几乎对于所有不超过 x 的偶数 n，都有 (11) 式成立。

这样一来，对于任意给定的正数 A，区间 $\left(\frac{x}{2},\ x\right]$ 中的偶数 n，除了可能有

$$\ll \frac{x}{\log^A x}$$

个例外值外，恒有

$$|D_1(n,x)| > |D_2(n,x)|$$

成立。若以 $E_1(x)$ 表示区间 $\left(\frac{x}{2}, x\right]$ 中的非哥德巴赫数的个数，则由此立即推出

$$E_1(x) \ll \frac{x}{\log^A x}。$$

这样就推出了定理 4.3.8 成立。

定理 4.3.8 是利用圆法和维诺格拉朵夫思想给予证明的。当一个强有力的思想问世之后。数学家很快就会接受过来，从而大大推进对于偶数哥德巴赫猜想的研究。研究猜想一方面要创造新的方法，另一方面也应对科学发展有强烈的敏感性，把其它创造的新思想新方法移植到自己所研究的问题上来。这样才会给研究工作带来生气勃勃的新局面，做出具有重大意义的成果。

对于一个猜想得到一个较弱结果之后，再向较强的结果一步步逼近，这是解决猜想的又一个重要途径。

1972 年，文汉（Vanghan）证明了

定理 4.3.9 存在正常数 c，使

$$E(x) \ll x\exp(-c\sqrt{\log x})。 \tag{12}$$

1975 年，Montgomery 和 Vanghan 进一步改进了（12），得到

定理 4.3.10 存在一个可计算的绝对正常数 Δ，使得

$$E(x) \ll x^{1-\Delta}。$$

为了证明这一结果，几乎用到了 L 函数零点分布的全部知识，并且把大筛法应用于对圆法中基本区间的讨论。

1979 年，我国两名著名的数学家陈景润和潘承洞定出常数 $\Delta > 0.01$。这是目前对于例外集合 $E(x)$ 的阶的估计最好的结果。

（四）筛　　法

为了证明把一个偶数拆成两个奇素数之和，我们探讨与此问题有关的更加广泛的命题：

把一个偶数拆成两个数 a 与 b 之和，其中 a 是一个不超过 a 个素因子的数，b 是一个不超过 b 个素因子的数。

这样两个数称为殆素数，记作 $(a+b)$。哥德巴赫猜想就是要证明 $(1+1)$。通过逐步减少素因子的个数的办法来寻求解决猜想（A）的途径，筛法就成了一个强有力的工具。

筛法是寻求素数的一个古老的方法。这个方法是两千多年前古希腊学者爱拉托斯（Eratosthenes，大约公元前 230）所创造的，称爱拉托斯筛法。用此方法可造出不超过已知数 N 的素数，现在叙述如下：

写出数 1，2，…，N，在这一列数中第一个大于 1 的数是素数 2。从数列中划掉 2 以外的所有 2 的倍数。接着 2 的第一个没有被划掉的数是素数 3。从数列中划掉 3 以外的所有 3 的倍数。接着 3 的第一个没有被划掉的数是素数 5，这样继续下去，就得到不超过已知数 N 的所有素数。

这是一种原始筛法，随着数学的发展，筛法也得到了发展。什么是筛法？现在用数学的语言叙述如下：

由有限个且满足一定条件的整数组成的集合以 A 表之，满足一定条件的无限多个不同的素数组成的集合记为 B，$z\geqslant 2$ 为任一正数。令

$$p(z)=\prod_{\substack{p<z\\ p\in B}} p。$$

在集合 A 中，所有与 $p(z)$ 互素的元素的个数记为 $S(A;B,z)$，即

$$S(A;B,z) = \sum_{\substack{a\in A \\ (a,p(z))=1}} 1 。$$

这里 $p(z)$ 就起到一个“筛子”的作用，凡是和它不互素的都被“筛掉”，而与它互素的数都被留下。所谓“筛法”其含义也正是如此。“筛子”的大小是与集合 B 及 z 有关。z 愈大，筛子就愈大，被筛掉的数就越多。$S(A;B,z)$ 是集合 A 经过筛子 $p(z)$ 筛选后所剩下的元素的个数。我们称 $S(A;B,z)$ 为筛函数。显然，筛法的目的就在于对于筛函数要了如指掌。因此研究筛函数的性质及其作用就成为“筛法”中的基本问题，而其中最重要的问题之一就是估计筛函数 $S(A;B,z)$ 的上界和正的下界。

设 A 是一由有限个整数组成的集合（元素可重复），B 是一个无限多个素数组成的集合。再设 $z\geqslant 2$ 是任意实数，并令

$$P(z) = \prod_{\substack{p<z \\ p\in B}} p。$$

易知筛函数具有如下简单性质：

i. $S(A;B,z) = |A|$①；

ii. $S(A;B,z) \geqslant 0$；

iii. $S(A;B,z_1) \geqslant S(A;B,z_2), 2\leqslant z_1 \leqslant z_2$；

iv.
$$S(A;B,z) = \sum_{a\in A} \sum_{d|(a,p(z))} \mu(d) = \sum_{d|p(z)} \mu(d)|A_d|, \qquad (1)$$

其中 A_d 表示集合A 中所有能被d 整除的元素所组成的子集合。

解决一个具体问题，就是归结到所给的问题如何与筛函数发生联系。现在把筛函数与命题 $(a+b)$ 的联系叙述如下。

设 N 为一大偶数，取集合

$$A = A(N) = \{n(N-n), 1\leqslant n \leqslant N\},$$

所有素数组成的集合记为 B。再设 $\lambda\geqslant 2$，取 $z = N^{1/\lambda}$。如果能

---

① $|A|$ 表示有限集合 A 的元素的个数。

证明筛函数

$$S(A;B,N^{1/\lambda})>0,$$

则显然就证明了命题 $(a+a)$，其中

$$a=\begin{cases}\lambda-1 & \lambda \text{ 是正整数；}\\ [\lambda] & \lambda \text{ 不是正整数。}\end{cases}$$

特别，当 $\lambda=2$ 时，这就证明了命题 $(1+1)$。

另一方面，若求得 $S(A;B,N^{1/\lambda})$ 的一个上界，那么我们就相应地得到一个大偶数表为二个素因子个数不超过 a 个数之和的表法个数的上界。

如果我们取集合

$$C=C(N)=\{N-p,p\leqslant N\},$$

能证明筛函数

$$S(C;B,N^{1/\lambda})>0,$$

则显然证明了命题 $(1+a)$。同样，若求得 $S(C;B,N^{1/\lambda})$ 的一个上界，那么，我们也就相应地得到了偶数表为一个素数与一个素因子不超过 a 个数之和的表法的上界。

由上述可知，命题 $(a+b)$ 和求筛函数的正下界与上界这一问题密切相连的。其中 z 不能取得太小（相对 N 来说），一定要取 $N^{1/\lambda}$那么大的阶。显然 λ 取得越小越好。如果一个筛法理论仅能对较小的 z（比如取 $\log N$）才能证明筛函数有正的下界估计，那么这种筛法理论对我们所讨论的问题是无用的。古老的筛法正是这样的。因此，要想解决我们的问题，必须发展已有的筛法。由 (1) 式可以看出，筛函数 $S(A;B,z)$ 的估计和集合 A_d，$d/p(z)$ 有关。如果对于给定的集合 A 及B，我们适当选取一个正数 $z>1$，及一非负可乘函数

$$\omega(d),\mu(d)\neq 0,(d,\overline{B})=1^{①},$$

---

① $\overline{B}$ 表示所有不属于B的素数组成的集合。设 μ 是一整数集合，d 为一整数，$(d,\mu)=1$ 表示 d 和 μ 中每一个数都互素。

并设

$$r_a = |A_d| - \frac{\omega(d)}{d}X, \tag{2}$$

我们的目的就是用 $\frac{\omega(d)}{d}X$ 代替 $|A_d|$。我们要求就某种平均意义上来说，使误差项 r_d 尽可能地小。怎样选取最好的 X 和 $\omega(d)$，这由集合 A 的性质来确定。

由 (1) 及(2) 有

$$S(A;B,z) = \sum_{d|p(z)} \mu(d)\frac{\omega(d)}{d}X + \sum_{d|p(z)} \mu(d)rd$$
$$= X\prod_{\substack{p<X \\ p\in B}}\left(1-\frac{\omega(p)}{p}\right) + \theta\sum_{d|p(z)} |r_d|, \quad |\theta| \leqslant 1。$$

当 z 相对于 X 并不是很大时，余项的项数 $\sum_{d|p(z)} 1$，即 $p(z)$ 的除数个数就可能很大，例如取 $p(z) = \prod_{p<z} p$，则当 $z > \log X$ 时，余项的项数就大于 X，这样就不可能得到有用的估计。这种方法仅当 z 很小时，例如 $z \ll \log\log X$，才有效。这就是所说的爱拉托斯筛法。这种筛法在理论上是无用的，因为数论问题所需要的是 z 相对于 X 来说是较大的情况。于是在 1920 年前后，布伦 (Brun) 首先对爱拉托斯筛法作了重大改进。布伦利用他的方法证明了命题 (9+9)。由于这一方法获得了对于哥德巴赫猜想研究的重大成果，这就开辟了人们利用筛法研究猜想 (A) 及其他数论问题的新途径。这种方法叫布伦筛法。1950 年前后，谢尔白格 (A. Selberg) 对爱拉托斯筛法，利用求二次型极值的方法，作了另一个重大改进。这种方法叫做谢尔白格方法。用该种方法，得到了筛函数的上界估计。这两种方法共同点在于设法控制余项的项数，使从余项所得的估计相对立项来说可以忽略不计，同时也要使主项得到尽可能好的估计。

把命题 $(a+b)$ 和对一个筛函数的估计直接相联系，这样得到的结果是较弱的。要得到较强结果，还要设法通过另一途径

来改进筛法。1941 年，库恩（Kuhn）首先提出了所谓“加权筛法”。后来数学家对各种形式的“加权筛法”进行了研究，从而使筛法的效用越来越大，所获得的结果也就得到不断的推进。

证明命题 $(a+b)$ 的历史进展可概述如下：

1920 年，布伦证明了命题 (9 + 9)；

1924 年，拉德马哈尔证明了命题 (7 + 7)；

1932 年，爱斯斯尔曼证明了命题 (6 + 6)；

1937 年，Ricci 证明了命题 (5 + 7), (4 + 9), (3 + 15) 以及(2 + 366)；

1938 年，布赫斯塔勃证明了命题 (5 + 5)；

1939 年，塔鲁塔柯夫斯基及 1940 年，布赫斯塔勃都证明了命题 (4 + 4)；

1941 年，库恩提出了“加权筛法”，后来证明了命题 $(a+b)$，其中 $a+b\leqslant 6$。

以上的结果都是利用 Brun 筛法得到的。以下的结果都是利用 Selberg 筛法得到的。

1956 年，王元证明了命题 (3 + 4)；

1957 年，维诺格拉朵夫证明了命题 (3 + 3)；

1957 年，王元证明了命题 (2 + 3) 以及命题 $(a+b)$，其中 $a+b\leqslant 5$。

为了证明命题 $(1+b)$，需要估计筛函数 $S(B;P,z)$。当估计筛函数的上界与下界时，需要对主要项进行计算，对余项进行估计。但在余项的估计上存在很大困难。这实质上，就归结到估计下面的和式：

$$R(x,\eta)=\sum_{d\leqslant x\eta}\mu^2(d)\max_{y\leqslant x}\max_{(l,d)=1}\left|\psi(y;d,l)-\frac{y}{\phi(d)}\right|。$$

对于这一和式进行估计，需要利用复杂的解析数论方法。

1948 年，匈牙利数学家 A. Renyi 利用列尼克所创造的大筛法，研究了 L 函数的零点分布，从而证明了：一定存在一个正

数 $\eta_0>0$，使对任意一个正数 $\eta<\eta_0$ 及任意正数 A，有估计式

$$R(x,\eta)\ll\frac{x}{\log^A x} \qquad (*)$$

成立。进而利用布伦筛法和这一结果证明了 $(1+b)$。

利用上述方法确定常数 η_0，将是很小的，而 b 将是很大的。我们希望 b 越小越好，这就需要改进方法，以便定出尽可能大的 η_0。

1962 年，潘承洞证明了当 $\eta_0=\frac{1}{3}$ 时，上面的估计式 (*) 成立，从而证明了 $(1+5)$。

1962 年，王元从进一步改进筛法着手，由 $\eta_0=\frac{1}{3}$ 推出了命题 $(1+4)$。同时还推得 η_0 和 b 间的一个非显然联系，从而分别推出命题 $(1+4)$ 和 $(1+3)$。

1962 年，潘承洞证明了 $\eta_0=\frac{3}{8}$ 时，估计式 (*) 成立，并利用较简单的筛法证明了命题 $(1+4)$。

1965 年布赫斯塔勃由 $\eta_0=\frac{3}{8}$ 推出了命题 $(1+3)$。

1966 年，陈景润宣布他证明了命题 $(1+2)$，1973 年，他给出该命题的详细证明。陈景润之所以能使哥德巴赫猜想研究推进一大步，是由于他提出了新的加权函数。对于同一个问题，选取不同的权函数，就可以得到不同的结果。当权函数 $\rho(a)=1$ 时，可得到命题 $(1+4)$，取 $\rho(a)=1-\frac{1}{2}\rho_1(a)$，就得到命题 $(1+3)$，而取

$$\rho(a)=1-\frac{1}{2}\rho_1(a)-\frac{1}{2}\rho_2(a),$$

就证明了命题 $(1+2)$。由此，我们可猜想，是否可用选取不同的权函数，去证明命题 $(1+1)$ 呢？可是按此方向考虑问题是否能走通，到目前为止，还看不出有什么眉目。

通过命题 $(a+b)$ 研究过程的简单概述，使我们看到，要推

进对猜想研究的结果，应在对已取得成果的基础上，对所用的方法作些不同方向上的改进和突破。方法的改进和突破是在猜想的研究中产生的。方法和成果是相辅相成的，因此，我们对于猜想的研究，应从不同的角度加以探索，这样不但有利于猜想本身的解决，而且在解决猜想的过程中还可以大大丰富数学内容，促使数学理论的发展。

五、补天何须五色石

——地图着色与四色猜想

在地图染色的实际工作中，提出了一个著名的数学难题——四色猜想。这一问题提出后，数学家经过一百多年的研究与探索，于1976年，借助于电子计算机才给出了证明，从而使四色猜想转化成四色定理。这是一个轰动整个数学界的重大事件。从思想方法的角度对这一事件作一分析，无疑是有重要价值和启发意义的。

（一）四色猜想的提出

1. 什么叫四色猜想？

画一张表现许多国家的地图，为了把不同的国家区别开来，常用不同的颜色来着色。在一张地图上，有多少国家，就用多少种颜色着色，肯定可以区别出来。但我们希望用的颜色种类要少，又要把国家之间明显地区分出来，即要保证相邻国家不用同一种颜色。这里所说的两个国家相邻，是指这两个国家之间有一条公共的国境线，否则是不相邻的。同时，这里所说的国家是连通一片的，不包括一个国家分成几个区域的情况。

作了上述说明之后，现在要问，画一张彩色地图，不管有多

少国家，也不管这些国家处在怎样的地理位置上，最少需要几种颜色？

经过大量的试验，画任何一张彩色地图，只要四种颜色就足够了。这就是地图四色问题，亦称“四色猜想”。这个问题的提法很简单，甚至连不识字的老太太也能听懂，但要证明这一猜测，一百多年来却绞尽了许多数学家的脑汁。直到 1976 年，才由美国数学家用电子计算机证明了它。从而四色问题就转化成了四色定理。不借助于电子计算机，至今尚无人用纯粹的数学手法给出其证明。

2. 先生问学生和学生问先生

1840 年，默比乌斯（A. F. Möbius, 1790—1868）给学生讲，在平面上很容易指出四个区域，其中每两个区域都有一个公共的边界线，并要求学生证明：在平面上决不可能指出五个区域都具有上述性质。从这个论断的证明中，可得出默比乌斯假设：画在平面或球面上的每张地图都可以用四种颜色着色，使得有公共边界的每两个国家都可用不同的颜色来着色。第一个明确提出四色问题的是一位大学生，名叫弗兰西斯·古特里。1852 年，弗兰西斯和他哥哥弗雷德里克同在伦敦上学。他写信给哥哥，指出每张地图上的国家总能用四种颜色着色，使相邻国家的颜色都不相同。这一问题能否用数学方法证明？哥哥回答不了弟弟提出的问题，就去问他们的老师、著名数学家 A. 德摩根（A. de Morgan, 1806—1871）教授。但 A. 德摩根也无法判定这一猜想的真伪性。于是，他又写信给他在三一学院的好友、著名数学家和物理学家哈密顿（W. K. Hamilton, 1805—1865）。他在信中写道：“我的一个学生今天要我为他提供一个充分的理由，来说明一件我自己还无法判明究竟是对还是错的事实。他说，如果画一张图，图上任意分成许多部分，凡是有共同边界线的两个部分都要涂上不同的颜色，那么，大概需要四种颜色，而不需要更多

的颜色就可以了。请问：难道不能构造出一个需要五种或者更多种颜色的图吗？”

哈密顿对这一问题也没有给出确切的回答。从这封信可以看出，当时德摩根对四色猜想是持怀疑态度的。经过探索，他证明了：五个国家不能每个都和其余的相邻。这一结果又使他相信永远不需要五种颜色，因而四色猜想是对的。但是，地图上不能有五个彼此相邻的国家的论据并不能成为四色猜想的证明。不然的话，如果六个国家中没有四个国家是每个都和其他三个相邻，就不需要四种颜色着色了。然而，实际上仍然需要四种颜色着色。所以，用地图所需色数等于相邻国家的最大值来推证四色问题是不行的，要解决四色问题需要另寻途径。

（二）早期的证明和五色定理

1. 凯利的呼吁

1878年，英国著名数学家凯利（Arthur Cayle, 1821—1895）在伦敦数学家会议上就四色问题作了报告。在报告中，他呼吁与会者独立地解决这一问题。于是，四色问题开始吸引了许多有才智的人去研究它，从而也就引起了人们的广泛重视。自凯利报告之后，各国所有数学中心和全世界所有主要数学杂志都源源不断地收到关于四色问题的种种错误证明。

凯利报告后的第一年，也就是1879年，发表了律师出身的A·B·肯普（A.B. Kempe, 1849—1922）对四色问题的证明。1880年，发表了P. J. 台特的证明。凯利和其他数学家对有关的论证没有发现什么破绽，都十分满意。但是在1890年，当时才29岁的年轻英国数学家P·J·希伍德（P.J. Heawood, 1861—1955）发现他们的证明是有漏洞的。后来有许多人设法弥补其漏洞，结果都失败了。虽然如此，肯普的文章仍然有价值，因为在他的证明

中包含了引导到正确证明的绝大部分基本概念，其中的思考路线可以用来获得新的成果。成果之一是可以用肯普的思考路线来证明五色定理。

2. 另辟蹊径

希伍德以其毕生精力来研究四色问题。他虽然没有最后解决这一问题，但他发表过好几篇重要的论文。他证明了任何地图只要五种颜色就够了，也就是地图的五色定理，或者叫希伍德定理。更重要的是希伍德想从考察一般曲面的着色问题来进攻平面四色问题。希伍德研究了更复杂的曲面上的地图着色问题。他证明了每张环面地图都可以用七种颜色着色，以使任何两个相邻国家都不会染上同一颜色。如果他的方法能用于平面，那就会提供四色猜想的证明。可惜这种解决方法只适用于更复杂的曲面，而不适用于简单的平面或环面。这表明一定的方法具有一定的适用范围，方法的运用是有条件的。

环面是只有一个“洞”的封闭曲面，由少到多，对于多个“洞”的情况，例如有四个“洞”的封闭曲面，又有怎样的结论呢？一般地，希伍德推测：

在有 $p \geqslant 1$ 个洞的封闭曲面上，足以为任何地图着色的最小色数等于

$$M_p = \left[\frac{7 + \sqrt{1 + 48p}}{2}\right],$$

其中 $[x]$ 表示取 x 的整数部分。特别是，当 $p = 1$ 时，$M_1 = 7$。这正是环面的结论。

后来，德国数学家 G. 林格尔用这个推测，首先证明了：足以为任何一张有 $p \geqslant 1$ 个“洞”的封闭曲面地图着色的真正最小色数 N_p，永远也不会超过“希伍德数” M_p，且有 $|N_p - M_p| \leqslant 2$。以后，美国数学家 U·T·杨斯进一步证明了 $|N_p - M_p| \leqslant 1$，而希伍德的假设对于不同于球面的“几乎一切”封闭曲面都是成

立的。最后，于 1974 年，林格尔对希伍德的假设作了完整的证明。

下面我们给出五色定理的证明，为此先介绍一些必要的知识。

3. 约当曲线和欧拉定理

我们假定，在平面上，有有限个弧，将平面划分为有限个区域。不妨假定任何两个弧或无交点，或有交点。在有交点的情况下，或交于一个共同端点，或交于两个共同端点。由一些弧及其端点所构成的图形，叫做网络。

约当曲线定理 在平面上任何一简单闭合曲线 c，把平面划分为两个区域，一个是有界的，即内部，一个是无界的，即外部。二者都以 c 为边界。

证明 不妨假定 c 是由有限个线段所组成的闭合折线，我们选择一坐标 xoy，使之具有如下性质：

(1) c 不与 y 轴相交；

(2) c 中的线段都不平行于 x 轴。

于是，任一平行于 x 轴的直线与 c 交点的个数必定是有限个。

在平面上，任取一点 P，作一直线 $PQ /\!/ ox$，且与 y 轴相交于 Q。PQ 与 c 交点的个数可这样计算：假若 PQ 与 c 的一个公共点不是 c 中任一线段的端点，那么该点算一个交点。假若 PQ 与 c 的一个公共点是 c 中某两线段的公共端点，且两线段在 PQ 的两边，那么该点也算是一个交点。否则不算交点。

假定 P 是平面上不在 c 上的点，如上法作 PQ，称 P 为第一类点；假若 PQ 与 c 交点的个数为奇数，称 P 为第二类点。假若其交点的个数为偶数，当 PQ 平行于 x 轴移动时，易见第一类点可以用不跨过 c 的线段结合在一起，第二类点亦然，但无法使第一类点与第二类点结合在一起。所以第一、第二类点分别形成

了两个区域。因为 y 轴上的点是属第二类的，故其形成的区域是无界的，而第一类点所形成的区域是有界的。

定理 若平面上有一网络 G，将其划分为有限个区域，则区域个数等于 1 的一个充要条件是 G 中不存在任何简单闭合曲线。

证明 必要性。若 G 中包含一个简单闭合曲线 c，则由上述定理得知 c 将平面划分为两个区域，故 G 将平面至少划分为两个区域。

充分性。若 G 中不包含任何简单闭合曲线，将 G 中弧的个数记为 n。我们用归纳法证明。当 $n=1$ 时，易见区域个数为 1。假定弧的个数少于 n，定理真，今推证当弧的个数等于 n 时亦真。事实上，由于 G 不包含任何简单闭合曲线，我们由 G 中一端点出发而沿 G 中的弧前进时，决不会再经过已到过的点，故最终到达一端点，再也无法前进。这就是说，G 中有一端点 P，它只是 G 中一个弧 E 的端点。我们去掉弧 E 和端点 P，得到新的网络记为 G'。在 G' 中有 $n-1$ 条弧，由假定，其区域的个数为 1。另一方面，平面被 G 划分所得的区域个数也是 G' 划分所得区域的个数，于是区域的个数仍然是 1。

欧拉定理 若平面有一连接的网络 G，其中至少有一端点，则 $F-E+V=2$，其中 F 表示 G 中区域的个数，E 表示弧的个数，V 表示端点的个数。

证明 用数学归纳法证明。先考虑 $F=1$ 的情况。这时对 E 用数学归纳法。当 $E=0$ 时，由假定知 $V=1$，于是有

$$F-E+V=1-0+1=2。$$

如果 $E=k(k\geqslant 0)$ 定理对，现证 $E=k+1$ 时，定理所述亦真。事实上，设 G 是连接网络，其中 $F=1$，$E=k+1$。由上述定理知 G 中不包含任何简单闭合曲线，于是 G 中存在一端点 P，它只是 G 中一个弧 c 的端点。去掉弧 c 及端点 P 得到一个新网络 G'。G' 中的区域、弧、端点的个数分别记为 F'，E'，V'，则有

$$F'=F=1,\ E'=E-1=k,\ V'=V-1。$$

由假定知

$$F'-E'+V'=2,$$

于是，对于 G 有

$$F-E+V=F'-E'+V'=2。$$

从而证明了当 $F=1$ 时，定理是对的。

如果假定 $F=i(i\geqslant 1)$ 时，定理真，能推出 $F=i+1$ 时，定理也对，则全部命题得证。事实上，设 G 中 $F=i+1$，由上面定理知，其中必存在一简单闭合曲线 c。若 ε 是 c 上一个弧，则在 G 中去掉 ε 得到 G'。G' 中的区域、弧及端点的个数分别记为 F'，E'，V'。由约当曲线定理知，ε 的两边是不同的区域。去掉 ε 后就会并成一个区域，于是 $F'=F-1=i$，$E'=E-1$，$V'=V$。所以，当 $F=i+1$ 时，有

$$F-E+V=F'-E'+V'=2。$$

4. 五色定理

五色定理　若平面上有一连通网络 G，由有限个弧及其端点所构成，将平面划分为有限个区域，则我们只需用五种颜色予以着色，使任何两个相邻区域有两种不同的颜色。

为了证明定理，我们先作三点说明。

(1) 不妨假定 G 中任何一个弧的两边是两个不同的区域。若不然，G 中有一弧的两边是同一个区域，我们可以从 G 中将这条弧去掉，这样去掉之后，既不影响区域的个数，也不影响两区域是否相邻的性质。

(2) 不妨假定 G 中每一端点是不多于三个弧的端点。如若不然，G 中 P 为 n 个弧的端点，则我们添加一个 n 边形的区域，使 G 变成 G'。在 G' 中每一个新端点有三条弧。

如果 G' 只需要五种颜色可着色，那么 G 也是这样，这只要让原有的区域保持已着色的颜色不变即可。

（3）如果 G 中有两个弧 ε_1 和 ε_2 仅有一个公共端点 P，且 P 点不是第三个弧的端点，我们就把 ε_1 和 ε_2 合并成一条弧 ε，其端点即 ε_1 与 ε_2 异于 P 的端点。这样改变显然不影响区域数和着色问题，所以我们可以一直作下去，直到不能再作为止。

经过上述改变之后，G 中每一端点是两个或三个弧的公共端点。若 P 只是两个弧 ε_1 和 ε_2 的公共端点，则 ε_1 和 ε_2 有第二个公共端点 Q。因为通过 Q 至多还有一个弧 ε，在 ε_1 与 ε_2 形成简单闭合曲线 c 的内部或外部。于是由于 G 的连接性质可知道 c 的外部或内部是区域之一。

由上面的说明之后，我们知道任何一个区域的边界是一简单闭合曲线，且由有限个弧构成。在 G 中，设其边界为 k 个弧所构成的区域的个数为 F_k，则存在一正整数 n，使 $2 \leqslant k \leqslant n$。

现在证 F_i，$i=2$，3，4，5 中至少有一个不为 0。如果在 G 中存在一端点 P，只是两个弧的公共端点，上面已证 $F_2 \neq 0$。现在考虑每一端点都是三个弧的公共端点。

设 F，E，V 分别表示 G 中区域、弧和端点的个数。

显然有

$$F = F_2 + F_3 + \cdots + F_n, \tag{1}$$

F_k 含 k 个弧。由于每个弧计算过二次，所以有

$$2E = 2F_2 + 3F_3 + \cdots + nF_n。 \tag{2}$$

因为每一端点有三个弧，所以

$$2E = 3V。 \tag{3}$$

由欧拉定理，得到

$$6F - 6E + 6V = 12,$$

由 (3)，上式可简化为，

$$6F = 3V + 12。$$

再由 (1)、(2) 得

$$4F_2 + 3F_3 + 2F_4 + F_5 = 12 + F_7 + 2F_8 + \cdots + (n-6)F_n \tag{4}$$

由此可见，在 G 中 F_i，$i=2$，3，4，5 至少有一个不为 0。如若不然，$F_2=0$，同样 $F_3=F_4=F_5=0$，因此，（4）式左边是 0，而右边至少是 12，这是不可能的。

五色定理是四色定理的减弱命题。当一个猜想一时难于解决的时候，往往考虑它的减弱命题。先解决较易解决的减弱命题，这一方面可以丰富数学的内容，另一方面也将为最终解决猜想积累经验、创造条件。这种解决问题的途径其实不是别的，它正是人的思维不断拓深的一种具体表现。正因其具有这样内在的基础，“归复杂为简单、尔后再服务于复杂问题的解决”的“迂回战术”才具有相当的普遍意义。

5. 肯普的证明

在一张地图上，如果满足下面两个条件：

（1）没有一个国家包围其它国家；

（2）没有三个以上的国家相遇一点。

那么，这张地图，就叫做是正规的。由上一节可知，对于一个非正规地图，经过适当修改后，可变成正规地图，且不影响着色及其相邻情况。因此，如果有一张需要五种颜色的地图，称为五色地图，那么就应该存在着一张五色正规地图。所以要想证明四色猜测，只要证明一张正规的五色地图是不可能就行了。由上一节可知，凡正规地图一定有一国具有少于六个的邻国。肯普（Kempe，Arithur Bray，英国人，1849—1922）注意到如果有一张正规的五色地图，就会有一张国数最少的地图，即“极小正规五色地图”。于是他提出：如果极小正规五色地图有一个国家的邻国数少于六个就会有一张国数较少的正规地图仍为五色的。所以也就不会有一个数恰好是极小五色地图的国数，因此没有极小五色地图的可能。肯普对四色猜测证明的关键是极小五色正规地图不能含有少于六个邻国的国家。由于肯普知道每张正规地图一定含有这样的一个国家，所以他断定没有需要五种颜色的极小正

规地图。因而不会有需用五种颜色的任何地图。下面我们就来详细考察他对有三个或四个邻国的国家是怎样进行证明的。

假设一张极小五色地图有一国 D 恰好有三个邻国 A，B，C。去掉 C，D 的边界线，则后者的国家数比前者少，因此可用四种颜色着色。若前者除了 D 外，都着上后者上的颜色，则 D 可着以与其三邻国不相同的颜色。这样一来，前者必然已用四种颜色着好色。这和五色的假设矛盾。

设一张极小五色地图有一国 E，恰有四个邻国 A，B，C，D。在该图中去掉 B，E 之间的边界线，使之合并成一个国家。这时可用四种颜色着色，剩下的合并国 E 未着色，如果 E 的四个邻国是用少于四种颜色着色，那么有一种颜色可供选择，供余下的国家用。这是可能的。这是因为在未着色的国家 E 的两旁的一对国家的颜色，例如 A（红），C（蓝），从其中一国到另一国要么有一条用这两种颜色着色的邻国的通道，要么没有这条通道。从 A 到 B 这条通道为 $AFGHIC$，而从 B 到 C 就不存在这样的通道。如果存在的话，那么一定有一个国家是两条通道所共有的，这当然是不可能的。因此两对中，必然有一对不存在这条通道。假设这对国家是 B 和 D，我们选择其中的一国 B，把所用的两种选择的颜色黄灰相继的国家一一列出来，这些国家是 B，U，V，W，其颜色次序为黄灰黄灰。我们交换一下次序为灰黄灰黄。于是未染色的国家 E 的邻国只有三色的邻国，因此 E 可用黄色着色。这样一来，这张地图就只要用四色就着色了。这与地图需要五色的假设相矛盾。

（三）四色猜想的证明

1. 不可避免组和可约构形

肯普曾证明每一张正规地图中至少有一国具有两个、三个、

四个或五个邻国，不存在每个国家多于五个邻国的平面正规地图。这用语句来表达即为：由有两个邻国、三个邻国、四个邻国及五个邻国组成的一组构形是不可避免的，即每张正规地图至少必须含有这四种构形中的一个。不可避免性是证明四色猜想的重要概念之一。另一个重要概念是可约性。直观上说，构形可约的条件是只要检查构形和成串国家能够合并的方式就有一种方法证明这个构形不可能出现在极小五色地图里。证明构形可约的方法是出自肯普证明有四个邻国的国家不可能出现在极小五色地图里。可约这个词来自肯普的论证形式：他证明只要五色地图中有一国具有四个邻国，就有国数减少的五色地图。

自从肯普第一次引入可约性概念以来，人们发展了检查构形是否可约的一些标准方法。使用这些方法来证明大的构形可约，需要检查大量的细节，似乎只有电子计算机才有可能做到。求出可约构形的不可避免组，那么四色猜想也就证明了。我们可以把肯普对四色猜想的证明看做是寻求可约构形的不可避免组的最初尝试。

2. 公开宣称的一种信念

1913 年，美国的数学家伯克霍夫（Birkhoff，George David，1884—1944）检查肯普证明的漏洞，发展了而后证明四色猜想的大半基础。伯克霍夫用肯普的想法和他自己的新技巧，能够证明某些大的构形可约。但直到 1950 年为止，证明了少于 36 国的地图可用四种颜色着色。

H. 希什（Heesch，Heinrich）1936 年开始研究四色猜想，似乎他是自肯普以后第一位数学家公开宣称一种信念，即四色猜想可用寻找可约构形的不可避免组来证明。1950 年，他猜想不仅能找到这样的组，而且这个组里构形大小估计为大约一万。这在当时要产生这样一个组并证明它的每个单元都是可约的似乎是不可能的。然而随着电子计算机的高速发展，为这个问题在技术

上提供了可能性。希什把证明构形可约的已知方法形式化，看出至少其中之一——把肯普使用的方法直接推广，在原则上是计算机所能做到的纯机械过程。

他的学生设计了一个程序，用来证明构形可约，有时成功，有时失败。但希什成功了：他能用这个程序产生的数据来证明构形可约，并且进一步的计算完成了一个更强有力的技术，它在原理上是属于伯克霍夫的。

由上面的叙述可知，一个对解决猜想有用的信念，需要经过许多数学家的探索与完善。有志于解决猜想的人，应在前人的基础上前进，才有可能走向成功的彼岸。当然，并非仅仅对猜想的解决依赖于历史，实际上人类的一切认识在一般的意义上讲都是历史和现实的统一——换言之，认识的完成或理论的发展都是在历史的基础上结合现实的客观分析努力探索的结果。

3. 等价的形式

希什描述可约性构形的方法比前人更加方便。他从改造原地图成为数学上称为“对偶图”入手，把原来的问题转化成与它等价的形式。其具体做法是：在地图上所有的国家内部选一点——相应国家的首都。我们约定，只有当两个国家有共同边界时，用一条铁路把两国连接起来，并让这条铁路通过这段边界。这时由首都（点）和连接它们的铁路（弧）组成的网构成一个图。该图称为原来地图上的图的对偶图。这样一来，把考虑地图上各国的着色问题，就转化为考虑它的对偶图上的顶点着色问题。

在认识事物的过程中，注意认识角度的变换是很重要的。其中较根本的一条途径就是排除认识中的干扰、消除与认识对象几不相干的因素，而这恰恰意味着对被认识对象的抽象或纯数学模型的提取。希什的改造体现的正是这种思想，这也就是其方法具有较前人进步性的主要原因，也是我们应加以继承和光大之处。

4．可约性障碍和放电

在正规地图中每个顶点恰好连接三条边。在这种情况下整个对偶图称为三角剖分。在其对偶图（连通的平面图）中，构形是三角剖分的一部分，由一组顶点加上连接这些点的所有边组成。边界上的圈称为构形的环，如六环构形，因为它的环有六个顶点。希什在检验构形的可约性时，观察到许多特殊现象可以提供约减成功的线索。例如，在有包括构形顶点的邻点在内的某些形式下从未找到可约构形。从未找到至少包括两个顶点的可约构形，其中一个顶点邻接四个环顶点，并且没有较小的可约构形。希什没有证明具有这些“约减障碍”的可约构形不存在，他发现了三个主要的约减障碍。希什引进了一个类似在电网络中移动电荷的方法——放电法，来求构形的不可避免组。这一方法虽然还很初步，但成为以后证明四色定理的关键环节。后来四色定理的证明，正是在希什工作的基础上完成的。

希什通过电网络移动电荷的方法，发现求构形的不可避免组的方法，充分说明了，解决数学问题的方法也受到人们对自然界解释的启发，特别是物理学可给我们提供线索。正如彭加勒（Henri Poincare，法国人，1854—1912）所说：“物理科学不仅给我们（数学家）以解决问题的机会，而且也帮助我们发现解决问题的方法，它把这贯穿于两种途径之中：引导我们去预测问题的解，以及启示适当论证的线索。”

5．新的困难

1970年，美国数学家黑肯（Haken，Wolfgang）试图以改良的放电过程的方法来证明四色猜想。但这里存在两大困难：第一，可约构形的任何一个不可避免组都可能会有很大的构形，其计算量大得惊人；第二，还不知道恰好需要多少个可约构形来形成一个不可避免组。在这时期，研究四色猜想的专家认为对于人

工用不太长的证明去攻克这个顽固堡垒似乎可能性很小。由于问题本身是那么浅显易懂，吸引成千上万的人努力去解决它。结果后来经过检查，这些“证明”都经不起推敲。虽然四色猜想没有得到证明，但为解决这一问题所运用的方法，却常常引导出数学领域中一些非常重要的结果。

1972年，黑肯根据当时所能利用的数学方法，作出如下的判断：肯定不会对四色猜想给出一个非机器证明。但在没有更为强有力的计算机之前能否用计算机给出证明也存有怀疑。问题的关键是要找出可约构形不可避免组问题的一组构形，其环点数少到足以使减约所需要的时间是在计算机所能完成的范围之内。为此，显然不必从检查全部构形的可约性入手。否则作估计所花的时间会超过全部工作的预期时间。我们只要把问题局限在某一定特点的问题上——“地理上的好”构形。这种构形其内部的顶点不能多于三个邻国在构形的环上，并且如果一个顶点正好有三个这样的邻国，那么这些邻国以相连的次序位于环上。

6. 人机合作证明了四色猜想

1972年，阿佩尔（Appel，Kenneth）和黑肯设计了一份计算程序，它能作出特殊类型的放电过程，它还能给出从最重要的情况得出的构形作为输出。经过计算机运行和不断修改程序，最后找到一个可行的程序，证明了地理上的好构形的不可避免组的存在性。但实际执行这个程序将会复杂到什么程度还知道得相当少。为此对于不含成对的相邻五次顶点的三角剖分的特殊问题进行了研究。就这样不断地实验、修改程序和改进放电过程，于1976年1月6日，阿佩尔、黑肯、科克确定了放电过程的细节。并由此得到了可约构形的不可避免组，设计了程序，用三部计算机，运行了一千二百多个小时，从而证明了四色猜想，解决了一百多年来人们想解决而一直没有解决的一个数学难题。其放电过程大约包括500个特殊放电情况，需要用人工分析约一万个顶点

带正电的邻近，需要用机器分析两千多个构形的可约性。

7. 解决地图四色问题的重大意义

地图四色问题的解决说明了有些问题不能单独用手工来完成，而必须靠人机结合才能给予解决。这就为数学的研究开辟了广阔的领域。这一生动的事例说明，只用逻辑推理的方法证明定理具有局限性，单靠计算方法解决问题也有一定局限性，而把这两种方法结合起来，就往往展现出一个新的天地。同时，在解决四色问题的过程中，也大大丰富了数学的内容和方法。图论中的许多重大成果都是直接或间接地在为解决四色问题过程中取得的，其中着色问题是直接为解决四色问题而展现出的重要课题。在攻克四色问题中又提出了一些新的猜想，新的矛盾又促使数学家用新的工具加以解决，数学因此在探索中得以不断前进。

（四）平面图

同一个图有各种不同画法，如果一个图能画在平面上，使得任何两条线除端点外是不相交的，那末这个图叫做可平面的图。

可平面的图 G 的 K-面着色是给 G 的所有面分配 K 种颜色，且使被一条边分离的两个面都不具有相同的颜色。使得 K-面可着色的那些 K 的最小值恰用 x^*（G）表示 G 的面色数。对于任何具有对偶 G^* 的可平面图 G，有

$$x^*(G) = x(G^*)。$$

与四色问题等价的命题有数十种，下面列举几种。

定理 5.4.1 任何平面图皆 4-可着色，当且仅当任何平面图皆 4-面可着色。

图 G 上的一个着色即一个映射 φ_S：$S \to M$，使得 $f(S_1) \neq f(S_2)$，当 S_1 与 S_2 不相邻，其中 $S = V, E, F$（若 G 为平面图），M 为一代数模，其上定义加法运算。记 $n(M)$ 表 M 中元素个数，且

规定次序。自然 $0\in M$。称这样的 ψ_S 为 G 对于 S 的色函数。M 称为色模。若存在 $s_0\in S$，使得 $\psi_s(s_0)=0$，则称 ψ_s 为正规的；若对任二色 α，$\beta\in M$，存在 s_1，$s_2\in S$ 相邻，且使得 $\psi_s(s_1)=\alpha, \psi_s(s_2)=\beta$，则称 ψ_s 是本原的。

一个边函数g：$E\to M$，将 E 中的边赋予方向，使得

$$g(e)=\begin{cases}|g(e)|, \text{当 } e=\langle u,v\rangle \text{ 为给定方向;}\\ -|g(e)|, \text{否则}\end{cases}$$

则称之为有向边函数。

定理 5.4.2 一个平面图 4-面可着色，当且仅当存在边集的 3-分解 $E=E_\alpha+E_\beta+E_\gamma$，使得对任意 $v\in V$，有

$$|E_u\cap E_\alpha|=|E_u\cap E_\beta|=|E_v\cap E_\gamma| \ (\mathrm{mod}2)\text{。}$$

定理 5.4.3 一个平面图 4-面可着色，当且仅当存在对于边的定向使得在任一圈 C 上，有

$$|E^+(c)|\geqslant\frac{1}{4}\varepsilon(c),\quad |E^-(c)|\geqslant\frac{1}{4}\varepsilon(c),$$

其中 $E^+(c), E^-(c)$ 分别表示 c 上对于c 的二个不同方向的边的集合。

在平面图 G 上，若 $H\subseteq G$，且 H 所有顶点都是偶次，则称 H 为 G 的欧拉子图。若 $f\in F$ (H)，且满足如下条件：

(1) 所有 $v\in\overline{V}\cap f$，都有 $|E_V\cap f|=0(\mathrm{mod}2)$，

(2) 对 $B(f)$ 的任意 连通片 L，有

$$\sum_{v\in\overline{V}(L)}|E_V\cap f|=0(\mathrm{mod}2),$$

则称 f 对于 G 全偶。

定理 5.4.4 一个平面图 G 为 4-面可着色，当且仅当存在一个欧拉子图 H，使得所有 $f\in F$ (H) 皆对 G 全偶。

从一个平面图的 4-面着色，可得边的 3-分解：$E=E_\alpha-E_\beta+E_\gamma$，其中 $E_t=\{e|g(e)=t\}=\{e|e$ 与 0，t 或$M-\{0,t\}$ 二色关联$\}$。

由此，可引入边函数

$$g(e)=\begin{cases}0, e\in E_\alpha,\\ 1, e\in E_\beta,\\ -1, e\in E_\gamma。\end{cases}$$

在任一顶点 V 处，由其旋所规定的任相邻边对形成一个角。记 $\Lambda=\langle c, e'\rangle$，$e$，$e'\in E_\alpha$，且 e' 继 e 之后。对于角，定义

$$\xi(\Lambda)=g(e')-g(e)(\text{mod}3),$$

称为角的示性函数。

定理 5.4.5 一个平面图为 4-面可着色，当且仅当存在一个角示性函数 ξ，使得 $\xi(\Lambda)=0$ 或 ± 1，且满足

(1) $\sum\limits_{\Lambda\subset B(f)}\xi(\Lambda)=0(\text{mod}3)$,

(2) $\sum\limits_{\Lambda\subset E_v}\xi(\Lambda)=0(\text{mod}3)$,

(3) $k_0=k_1=k_{-1}(\text{mod}3)$。

定理 5.4.6 所有平面图 4-面可着色，当且仅当所有 3-正则平面图 4-面可着色。

定理 5.4.7 3-正则平面图 4-面可着色，当且仅当极大平面图有偶对分解。

（五）线（边）着色

我们可以把三次平面图的面染色问题进一步转化为三次平面图的线着色问题。

定义 所谓地峡线是指这样的一条线，从 G 中去掉这条线之后，G 就分为两个无公共点的子图。

定理 5.5.1 设 G 是一个三次平面图且无地峡线，则 G 的线可以用 3 种颜色去染，使得任何两条相邻（有公共端点）的线具有不同的颜色，当且仅当 G 的面是四色可染的。

证明 若 G 的面是 4-色可染的，设这 4 种颜色为 A，B，C，D。我们构造一个四元群如下：

$\oplus$	A	B	C	D
A	D	C	B	A
B	C	D	A	B
C	B	A	D	C
D	A	B	C	D

由于 G 是三次的，且无地峡线，所以 G 的任何两个相邻的面有唯一的一条公共边界线，并且每一条线是唯一的两个相邻面的公共边界线。因此，我们可以按如下方式对 G 中的线进行染色：若线 e 是具有颜色 X 和具有颜色 Y 的两个面的公共边界线，则将 e 染为颜色 $X \oplus Y$。由于 $X \neq Y$，所以 $X \oplus Y \neq D$。这样一来，与每一点关联的三条线，恰好染三种不同的颜色 A，B 和 C，即 G 的线是三色可染的。

若 G 是线 3-色可染的，那么染 A，B，C 色的集合分别以 $f(A), f(B), f(C)$ 记之。显然，$f(A), f(B), f(C)$ 是 G 中线的一个分划。G 中由 $f(A) \cup f(B)$ 所组成的子图以 G_{AB} 记之。由于 G 为三次的，所以 G_{AB} 由一些没有公共点的圈组成，从而可用两种颜色 α，β 对 G_{AB} 的面进行染色，使得任何相邻的两个面具有不同的颜色。同样考虑图 G_{AC}，设用 r 和 o 两种颜色对 G_{AC} 的面进行染色。这样一来，G 的每一个面都染上了两种颜色，或 αr 色，或 αo 色，或 βr 色，或 βo 色。令 $A = \alpha r, B = \alpha o, C = \beta r, D = \beta o$。不难看出，$G$ 中任何相邻的两个面具有不同的颜色。

由此可见，平面图的四色问题，可以化为三次平面图的线的三着色问题。由图的 3-着色问题，一般地，可提出图的线的 k-着色问题。

对于任一平面图 $G = (V, E)$，所谓 G 对 E 的 k-着色，是指对 G 的如下分解

$$E = E_1 + E_2 + \cdots + E_k,$$

使得对于所有的 $i(1\leqslant i\leqslant k)$，任 $e,e'\in E_i$，均有 e,e'不相邻，即 E_i 皆独立集。若 G 有一线 K-着色，则称 G 为线 K-着色。记

$$x_E(G)=\min\{K\mid G\text{ 为线 }K\text{-着色}\},$$

称为 G 的区域色数。

一个既没有圈也没有两条边连接同一对顶点的图称为简单图。没有任何边的图称为是空图。一个图的顶点集若能分解为两个子集 X 和 Y，使每条边有一个端点在 X 中，另一个端点在 Y 中，则称此图为二分图。

图 H 是 G 的子图（记为 $H\subseteq G$），若 $V(H)\subseteq V(G)$，$E(H)\subseteq E(G)$。若 $H\subseteq G$，但 $H\neq G$ 时，则记为 $H\subset G$，这时称 H 为 G 的真子图。G 的顶点 V 的度数（记为 $d_G(V)$）是指 G 中与 V 关联的边的数目，每个圈算作两条边。以 $\delta(G)$，$\Delta(G)$ 分别表示 G 的顶点的最小度数、最大度数。$C(G)$ 表示 G 的闭色。

定理 5.5.2 当 G 为二分图时，则

$$x_E(G)=\Delta。$$

若证明此定理需要用到下面两个引理。

引理 1 设 G 是一个非奇回路的连通图，则 G 有一个 2-边着色，其中两种颜色在度数至少为 2 的每个顶点上都出现。

引理 2 设 $b=(E_1,E_2,\cdots,E_K)$ 是 G 的一个最优 K-边着色，若存在 G 中一个顶点 u 和颜色 i 及 j，使 i 不出现在 u 上，而 j 至少两次出现在 u 上，则 $G(E_i\cup E_j)$包含 u 的那个 Q 分支是一条奇回路。

现在我们证明定理。

显然，在任何正常线着色中，和任何一个顶点关联的各边必须分配以不同的颜色，因此，有

$$x_E(G)\geqslant\Delta。$$

现在证明对于二分图来说，大于号是不可能的。如若不然，即设二分图 G，有

$$x_E(G) > \Delta。$$

设 $b=(E_1,E_2,\cdots,E_\Delta)$ 是 G 的一个最优 Δ- 线着色，并设 u 是使 $c(u)<d(u)$ 的一个顶点。显然 u 满足引理 2 的假设，于是 G 包含一个奇回路，这与 G 假设为二分图相矛盾。从而定理证明完毕。

定理 5.5.3 若 G 是简单图，则 $x_E(G)=\Delta$ 或 $x_E(G)=\Delta+1$。

证明 由于对任意图皆有 $x_E(G)\geqslant\Delta$，因此只需要证明 $x_E(G)\leqslant\Delta+1$ 即可。为此，我们用反证法。假设 $x_E(G)>\Delta+1$。设 $v=(E_1,E_2,\cdots,E_{\Delta+1})$ 是 G 的一个最优 $(\Delta+1)$- 边着色，并设 u 是一个满足 $c(u)<d(u)$ 的顶点。则存在颜色 i_0,i_1，使 i_0 不出现在 u 上，而 i_1 至少两次出现在 u 上。设 uv_1 具有颜色 i_1，由于 $d(v_1)<\Delta+1$，又一颜色 i_2 不出现在 v_1 上。这样 i_2 必然出现在 u 上。因为如若不然，用 i_2 来给 uv_1 重新染色，可得 b 的改变。因此某边 uv_2 有颜色 i_2。继之，由于 $d(v_2)<\Delta+1$，某一颜色 i_3 不出现在 v_2 上，i_3 必然出现在 u 上，因否则的话，仿 i_2 给 uv_1，用 i_3 给 uv_2 重新染色，这样便仅得到一个改变的 $(\Delta+1)$- 边着色。所以某边 uv_3 有颜色 i_3。如此做下去，我们得到了一个顶点序列 v_1，v_2，…和颜色序列 i_1，i_2，…，使

(1) uv_j 有颜色 i_j。

(2) i_{j+1} 不出现在 v_j 上。

由于 u 的度数是有限的，故存在一个最小整数 l，使得对于某一个 $k<l$。

(3) $i_{l+1}=i_u$。

现在以如下方式给 G 重新染色。对于 $1\leqslant j\leqslant k-1$，用颜色 i_{j+1}，给 uv_j 重新染色，产生一个新的 $(\Delta+1)$- 边着色 $b'=(E_1{}',E_2{}',\cdots,E'_{\Delta+1})$。显然

$c'(v)\geqslant c(v)$，对所有的 $v\in V$ 成立。于是，b' 也是 G 的

最优$(\Delta+1)$-边着色。由引理2知，$G(E'_{i_0}\cup E'_{i_k})$含u的分支H'是一条奇回路。

继而，用颜色i_{j+1}给uv_i重新染色，$k\leqslant j\leqslant l-1$，并且用颜色$i_k$给$uv_l$重新染色，得到一个$(\Delta+1)$-边着色$b''=(E''_1, E''_2, \cdots, E''_{\Delta+1})$。于是，有

$c''(v)>c(v)$，对所有的$v\in V$，且$G(E''_{i_0}\cup E''_{i_k})$含$u$的分支$H''$是奇回路。但是，由于$v_k$在$H'$中有度数2，$v_k$在$H''$中有度数1。矛盾。故命题得证。

进一步，Vizing于1964年证明了更一般形式的定理如下：

定理5.5.4 若G是无圈的，则$\Delta\leqslant x_E(G)\leqslant\Delta+\mu$，其中$\mu$为$G$的重数，所谓$G$的重数，即在$G$中连接两个顶点的边的最大数目。

定义 一个简单图，若它的每一对不同的顶点均有一条边相连，则称它为一个完备图。

定理5.5.5 对于完备图$k_n(n\geqslant 2)$，

$$x_E(k_n)=\begin{cases}n,\text{当 } n \text{ 为奇数},\\ n-1,\text{当 } n \text{ 为偶数}。\end{cases}$$

定理5.5.6 设G^c为G的补图，n为G的顶点数，

当n为奇数时

$$n\leqslant x_E(G)+x_E(G^c)\leqslant 2n-3,$$

$$0\leqslant x_E(G)x_E(G^c)\leqslant(n-1)(n-2);$$

当n为偶数时

$$n-1\leqslant x_E(G)+x_E(G^c)\leqslant 2n-2,$$

$$0\leqslant x_E(G)x_E(G^c)\leqslant(n-1)^2。$$

定理5.5.7 设G是一个无环的多重图，其中$x_E(G)=k+1$，若除了边ab以外，用k种颜色染G的其他所有边。令C_x表示在顶点x处的颜色集合，则

$$|C_a\cup C_b|=k,$$

$$|C_a \cap C_b| = d_G(a) + d_G(b) - k - 2,$$
$$|C_a \setminus C_b| = k - d_G(b) + 1,$$
$$|C_b \setminus C_a| = k - d_G(a) + 1。$$

（六）顶点着色

上面已经谈到过平面图的面着色问题，可转化为平面图的顶点着色问题。下面给予确切的叙述。

给定一个平面图 $G=(V, E)$，构造另一个平面图 G^* 如下：在 G 中每一个面 f_i 内，取一点 x_i 作为 G 的点，若 G 中的线已是面 f_i 和 f_j 公共边界线，则在 x_i 和 x_j 之间连一条线 l，当 e 是地峡线时，f_j 和 f_i 是同一个面，这样就得到了图 G^*，称为 G 的对偶图。

显然，若 G 无地峡线，则当且仅当 G^* 中的点可以用 K 种颜色进行染色，使得任何相邻（有线相连）的两个点具有不同的颜色时，G 的面可以用 K 种颜色染色，使得任何相邻两个面具有不同的颜色。

因此平面图的面着色问题又等价于平面图的点着色问题。对于顶点着色，我们可以推广到一般的图上。

图 G 的一个 K-顶点着色问题是指用 K 种颜色着到 G 的各顶点的一种分配，且使相邻接的顶点没有染有相同的颜色。使 G 存在一个 K-顶点着色的最小的 K 称为 G 的顶点的色数，记为 $x_v(G)$。若 $x_v(G) = K$，G 称为是 K 色的。

当 G 有一个 K-顶点着色时，称 G 是 K-可着色的。显然一个简单图，当且仅当它是空图时，才是 1-可着色的；当且仅当它是二分图时，才是 2-可着色的。

1952 年，Dirac 提出临界图的概念。所谓一个图 G 是临界的是指对 G 内每个真子图 H，均有 $x_V(H) < x_V(G)$。一个 K-临界图是指既是 K-色的又是临界的图；每一个 K-色图有一个 K-临

界子图。对于这种特殊类型的临界图的研究，有助于处理着色问题。下面我们给出临界图的基本性质。

定理 5.6.1 若 G 是K-临界的，则 $\delta \geqslant K-1$。

证明 如若不然，设 $\delta < K-1$，再设 V 是C 中度数为δ 的一个顶点。由于 G 是K-临界的，所以 $G-V$ 是（$K-1$）-可着色的。由于 V 在G 中与δ 个顶点相邻（$\delta < K-1$），因此在 $K-1$ 种颜色中至少有一种颜色 i，使任何与 V 相邻接的点都与i 色不同。故令 V 点染i 色，于是得 G 为$K-1$ 着色。矛盾。故必有 $\delta \geqslant K-1$。

推论 1 任何 K-色图中至少有K 个点的度数不小于$K-1$。

证明 设 G 是K 色的，H 是G 的一个K-临界子图。$x_V(H) = x_V(G) = K$。由定理 5.6.1 知，H 的每个顶点在H 中的度数不小于$K-1$，因而在 G 中的度数也不小于$K-1$。由于 H 是K-色的，显然，它至少有 K 个顶点的度数不小于$K-1$。

推论 2 对任何图 G，有

$$x(G) \leqslant \Delta(G) + 1$$

证明 若 $x(G) \geqslant \Delta(G)+2$，由上可知，$G$ 中至少有$x(G)$个点度数不小于 $x(G)-1 \geqslant \Delta(G)+1$。因 $\Delta(G) > 0$，即 G 中存在至少为 $\Delta(G)+1$ 的点，与 $\Delta(G)$ 的定义矛盾。于是结论得证。

G 的顶点 V 称为是一个割点，若其边 E 可以分为两个非空子集E_1 和 E_2，使 $G[E_1]$ 和 $G[E_2]$ 正好以 V 作为公共顶点。没有割点的连通图称为块。

$V(G)$ 的一个子集 S，其中任何两点是不邻接的，则称 S 为G 的一个独立集。一个独立点集 S 称为是最大的，若对任一独立集 S'有：$|S'| \leqslant |S|$（其中 $|S|$ 表示集 S 的元素个数）。简单图 G 的一个团是指V 中的一个子集S，使 $G[S]$ 是完备的。

定理 5.6.2 在临界图中，没有顶点割是团。

证明 用反证法。设 G 是一个K-临界图，并假设 G 有一个顶点割 S 是一个团。记 G 的S-分支为G_1，G_2，…，G_n。由于

G 是K-临界的，所以每个 G_i 是$(K-1)$-可着色的。进一步，因为 S 是一个团，故 S 为完备图，所以 S 中的各顶点必然在G_i 的任何 $(K-1)$ -着色中染上相异的颜色，有 $|S|$ 种颜色。由此可知，存在 $G_1, G_2, \cdots, G_n$ 的一组$(K-1)$-着色，它们在 S 上符合。但是这些着色合在一起产生 G 的一个$(K-1)$-着色，矛盾。

推论 3 每个临界图都是一个块。

证明 若 V 是一个割点，则 $\{V\}$ 是一个顶点割，当然也是一个团。而从上述定理知，没有一个临界图具有割点；换言之，每个临界图是一个块。

推论 2 中给出的色数上界，有时比实际值大许多，那么这一结果是否可进一步加以改进呢？经探索对于某一类图这是可能的。下面我们给出 Brooks 定理。为此先引入奇回路的概念。若一条闭链的起点和内部顶点互不相同，则它称作是一回路。其长为 K 的回路，叫做 K-回路，特别当 K 为奇数时，就叫做奇回路。

定理 5.6.3 (Brooks 定理) 若 G 是连通的简单图，并且它既不是奇回路，又不是完备图，则

$$x(G) \leqslant \Delta。$$

证明 不失一般性，满足定理假设的 K-色图 G 是 K-临界的。由推论 3 知，G 是一个块。同时，由于 1-临界图与 2-临界图皆是完备的，而 3-临界图则是奇回路，故有 $K \geqslant 4$。

若 G 有一个 2-顶点割 $\{u, v\}$，则可推出

$$2\Delta \geqslant d(u)+d(v) \geqslant 3K-5 \geqslant 2K-1。$$

由于 2Δ 是偶数，这就蕴含着 $x(G)=K \leqslant \Delta$。

其次，假设 G 是 3-连通的。由于 G 不是完备的，所以在 G 中存在三个顶点 u，v 和 m，使 uv，$vw \in E$，而 $uw \overline{\in} E$。设 $u=v_1, w=v_2$，且 v_3，v_4，$\cdots$，$v_l=v$ 是$G-\{u, w\}$ 的顶点的任意一个排列，使每个 v_i 都和某一个具有$j>i$ 的 v_j 相邻接。现在给 G 的一个Δ-着色如下：在 $v_1=u$，$v_2=w$ 上染上颜色 1，

然后按颜色表 1，2，…，Δ 中最先可用的颜色依次给 v_3，v_4，…，v_l 染色。据 v_1，v_2，…，v_l 的构造，每个顶点 $v_i(1\leqslant i\leqslant l-1)$都和具有 $j>i$ 的某一个顶点 v_j 相邻接，也都和具有 $j<i$ 的最多和$\Delta-1$ 种颜色相邻接，于是颜色 1，2，…，Δ 中必有一种可以给 v_i 染上。最后，由于 v_l 和已染上颜色 1 的两个顶点 v_1 和 v_2 相邻接，因而它最多和另外的 $\Delta—2$ 种颜色相邻接，于是颜色 2，3，…，Δ 中必有一种颜色可以用来给 v_l 染上。定理证毕。

图 G 的一个剖分是指把 G 的边进行一系列剖分而得到的一个图。

K-色图的充要条件是什么？当 $K=1$，$K=2$，$K=3$ 时的必要条件容易回答。1952 年 Divarc 研究 $K=4$ 时的情况，于是 1961 年 Hajos 提出 K-色图的必要条件是：若 G 是 K-色的，则 G 包含 K_4 的一个剖分。这就是 Hajos 猜想。这个猜想至今还没有解决，这是被公认的一个难题。

在研究着色问题时，我们不仅要考察着色的存在性，而且还要把着色的具体数目算出来。用$\pi_K(G)$表示 G 的相异 K-着色的数目。当且仅当 G 是 K-可着色时，$\pi_K(G)>0$。如果在两个着色中某一顶点被分配以不同的颜色，则认为这两种着色是相异的。例如，有三顶点三边的图 3-着色，就有六个着色是相异的。

若 G 是空的，则每个顶点可以独立地分配以 K 种有用颜色中的任何一种。所以$\pi_K(G)=K^r$。若 G 是完备图，则第一个顶点可以在 K 种颜色中选择一种，对第二个顶点可以在 $K-1$ 种颜色中选择一种，如此等等。因此，在这种情况下，有$\pi_K(G)=K(K-1)\cdots(K-v+1)$。一般地，有

定理 5.6.4 若 G 是简单图，则对 G 的任何边 e

$$\pi_K(G)=\pi_K(G-e)-\pi_K(G\cdot e)$$

均成立。

证明 设 e 的两个端点为 u，v。对于在 u 和 v 上分配以相同颜色的 $G\text{-}e$ 的每一 K-着色，对应着 $G\cdot e$ 的一个 K-着色，其中由 u 和 v 等同起来所形成的 $G\cdot v$ 的那个顶点被分配以和 u 与 v 相同的颜色。这个对应显然是可逆的一一满对应。所以 $\pi_K(G\cdot e)$ 恰好是使 u 和 v 分配有相同颜色的 $G-e$ 的 K-着色的数目。

同样，由于使 u 和 v 分配有不同颜色的 $G\text{-}e$ 的每个 K-着色都是 G 的一个 K-着色。所以，反过来，π_K（G）是使 u 和 v 分配有不同颜色的 $G\text{-}e$ 的 K-着色的数目。由此可知：

$$\pi_K(G-e)=\pi_K(G)+\pi_K(G\cdot e),$$

等式移项，就得到定理所要求的结论。

推论 4 对于任何图 G，$\pi_K(G)$都具有整系数的关于 K 的 v 阶多项式，首项为 k^v，K^{v-1}的系数为 ε，常数项为 0。进一步 $\pi_K(G)$ 系数的符号是交替的。

证明 对 ε 进行归纳法。不妨假设 G 是简单图。若 $\varepsilon=0$，则 $\pi_K(G)=K^v$，显然满足推论条件。若边数少于 m 的所有图皆对，且设 G 具有 m 条边($m>1$)。设 $e\in G$，则 $G-e$ 和$G\cdot e$ 两者都有 $m-1$ 条边。由归纳假设知，存在非负整数 a_1，a_2，…，a_{v-1}和 b_1，b_2，…，b_{v-2}，使得

$$\pi_K(G-e)=\sum_{i=1}^{v-1}(-1)^{v-i}a_iK^i+K^v$$

和
$$\pi_K(G\cdot e)=\sum_{i=1}^{v-2}(-1)^{v-i-1}b_iK^i+K^{v-1}。$$

根据上述定理，有

$$\pi_K(G)=\pi_K(G-e)+\pi_K(G\cdot e)$$
$$=\sum_{i=1}^{v-2}(-1)^{v-i}(a_i+b_i)K^i-(a_{v-1}+1)K^{v-1}+K^v。$$

于是 G 也满足推论条件，根据归纳原则得所欲证。

1943 年 Hadwiger 提出猜想：若 G 是K-色的，则 G 可“收缩”为一个包含 K_K 的图。1964 年 Wagher 证明了 $K=5$ 时，

Hadwiger 猜想等价于著名的四色猜想。

（七）全色猜想

由上面的讨论中可看出在解决四色问题的过程中，研究其等价命题，将其内容扩展到一个更加广阔的范围，使之讨论研究的内容越来越丰富，并且还把内容扩展到数论和方程的范围之中，同时还提出了很多有待解决的问题。例如，对于 $m < n$，设 $f(m,h)$表示不包含 K_n，但在每个 2-边着色中都存在一个单色 K_m 的图的最少可能的顶点数。1970 年 Folkman 给出这个图的存在性，确定$f(m,n)$的界限。目前已经知道的结果为：

$$f(3,n)=6, \text{对于 } n \geqslant 7,$$

$$f(3,6)=8,$$

$$10 \leqslant f(3,5) \leqslant 18,$$

其一般情况还有待于研究。

上面我们对一个图的着色问题，分别考虑了边着色、点着色问题。那么这两者之间又有什么联系呢？下面我们就来讨论这个问题。

设 $G=(V, E)$ 是一个简单图，构造一个图 $\bar{G}$ 如下：$\bar{G}$ 中每一个点对应于G 中的一条线，当且仅当 G 中两条线相邻时，$\bar{G}$ 中与这两条线对应的两个点之间有线相连，我们称 $\bar{G}$ 为G 的线图。

显然，当且仅当 G 中的线可以用K 种颜色去染，使得相邻的两条线具有不同的颜色时，$\bar{G}$ 中的点可以用K 种颜色去染，使得任何相邻两点有不同的颜色，因此一个图的线着色问题，可以化为它的线图的顶点着色问题。由此可见，如果图的顶点着色问题解决了，那么线着色和面着色问题也就自然解决了。

如果把边着色和顶点着色联系起来考虑又会怎样呢？即对于一个图 G 进行染色，使邻接的边或顶点不染同样的颜色，至少

需要几种颜色呢?

一般地，其色数是否为 $\Delta+2$ 呢? 1965 年，M. Behzad 经过大量的事例分析，提出了如下的“全色猜想”:

对于任意简单图 G，$V \cup E$ 的元素可以用 $\Delta+2$ 种颜色着色，使没有两个邻接的或关联的元素染有相同的颜色。

1971 年，M. Rosenfeld 和 N. Vijayaditya 对于 $\Delta \leqslant 3$ 的情况证实了这一猜想是对的，对于一般情况还有待人们去探索。

六、法无定法

——提出数学猜想的若干方法

前面我们系统地回顾和分析了在历史上出现的一些重要数学猜想。从这些回顾与分析中，不难看出，发现并提出数学猜想，不仅需要有雄厚的数学知识，而且还要有行之有效的科学方法。提出数学猜想的方法是多种多样的，其表现形式又是极其复杂的。这里，仅就其中几种主要的方法，作以初步分析。

（一）不完全归纳法

不完全归纳法是提出数学猜想的一种常见的方法。这种方法的基本思想是，根据某类数学对象中一些个别对象具有某种属性而猜测该类对象全体都具有这种属性。利用这种方法发现和提出猜想，有时是在某些具体计算基础上推想出来的。像前面讲过的哥德巴赫猜想，就是运用这种方法提出来的。他首先发现对于较小的自然数，把一个偶数拆成若干组两个奇数之和时，其中至少有一组是两个奇素数；把一个奇数拆成若干组三个奇数之和时，其中至少有一组均为奇素数。然后，他根据这些最初的有限验算，大胆提出了猜想：所有每个大于 4 的整数都可以表示为两个素数之和。这个猜想看起来并不复杂，但自 1742 年提出至今已两个多世纪了，仍未最后解决。下面我们再举出几个这方面的事

例。

1644 年，默森尼（Mersenne）研究并通过计算得知，当 P 为 2，3，5，7，13，17，19，31，127 时，形如 $M(P)=2^P-1$ 的数（亦称“默森尼数”）为：

$$M(2)=2^2-1=3,$$
$$M(3)=2^3-1=8-1=7,$$
$$M(5)=2^5-1=32-1=31,$$
$$M(7)=2^7-1=128-1=127,$$
$$M(13)=2^{13}-1=8192-1=8191,$$
$$M(17)=2^{17}-1=13072-1=131071,$$
$$M(19)=2^{19}-1=524288-1=524287,$$
$$M(31)=2^{31}-1=2147483648-1=2147483647,$$
$$M(127)=2^{127}-1$$
$$=170141183460469231731687303715884105728-1$$
$$=170141183460469231731687303715884105727,$$

这些数均为素数。据此，提出猜想：形如 $M(P)=2^P-1$（P 为素数）的数中有无限多个素数。到 1979 年为止，人们已发现有 27 个这样的素数，即当 $P=2$，3，5，7，13，17，19，31，61，89，107，127，521，607，1279，2203，2281，3217，4253，4423，9689，9941，11213，19937，21701，23209，44497 时，则 $M(P)=2^P-1$ 均为素数。

1664 年，法国数学家费尔马研究了形如 $F(n)=2^{2^n}+1$ 的数（$n\geqslant 0$ 的数），并且具体计算以下五个数：

$$F(0)=2^{2^0}+1=2+1=3,$$
$$F(1)=2^{2^1}+1=2^2+1=5,$$
$$F(2)=2^{2^2}+1=2^4+1=17,$$
$$F(3)=2^{2^3}+1=2^8+1=257,$$

$F(4) = 2^{2^4} + 1 = 2^{16} + 1 = 65537$。

由于上述五个数都是素数，故费尔马提出推断：对于任何自然数 n，形如 $F(n) = 2^{2^n} + 1$ 的数都是素数。这就是著名的费尔马猜想。

杰波夫（Desboves）猜想：相邻平方数之间至少存在二个素数，也是采用这种不完全归纳法提出来的。事实上，此猜想可以表述为："对于任何自然数 n，在 n^2 与 $(n+1)^2$ 之间至少存在二个素数"。通过一些具体的计算得：

当 $n=1$ 时，

$n^2=1$ 与 $(n+1)^2 = 4$ 之间有素数 2,3；

当 $n = 2$ 时，

$n^2 = 2^2 = 4$ 与 $(n+1)^2 = 3^2 = 9$ 之间有素数 5,7；

当 $n = 3$ 时，

$n^2 = 3^2 = 9$ 与 $(n+1)^2 = 4^2 = 16$ 之间有素数 11,13；

当 $n = 4$ 时，

$n^2 = 4^2 = 16$ 与 $(n+1)^2 = 5^2 = 25$ 之间有素数 17,19,23；

当 $n = 5$ 时，

$n^2 = 5^2 = 25$ 与 $(n+1)^2 = 6^2 = 36$ 之间有素数 29、31；

当 $n = 6$ 时，

$n^2 = 6^2 = 36$ 与 $(n+1)^2 = 7^2 = 49$ 之间有素数 37,41,43,47；

当 $n = 7$ 时，

$n^2 = 7^2 = 49$ 与 $(n+1)^2 = 8^2 = 64$ 之间有素数 53,59,61；

当 $n = 8$ 时，

$n^2 = 8^2 = 64$ 与 $(n+1)^2 = 9^2 = 81$ 之间有素数 67,71,73,79；

………………

根据上述结果作出推断："相邻平方数之间至少存在二个素数"。这一数学猜想与"当 n 为任何自然数时，在 n^2 与 n^2+2n+1 之间至少有二个素数"是等价的。

还有，数论中的所谓“凯特兰（Catalan）猜想”：除 $8=2^3$，$9=3^2$ 之外，没有两个连续整数都是正整数乘幂，也是通过一些具体的结果而作出的一种推断。

利用不完全归纳法提出数学猜想，不仅表现在通过一些个别计算结果作出一般判断，而且还表现在通过一些特殊推理作出普遍结论。比如，数学家波文（Bowen）研究方程：

$$1^n+2^n+\cdots+m^n=(m+1)^n。$$

发现该方程有解：$n=1$，$m=2$，但再未发现其他解，故提出此方程只有 $n=1$，$m=2$ 这组正整数解，即所谓“波文猜想”。这个猜想至今未解决。我国数学家柯召和孙琦研究了更为一般的方程：

$$x^n+(x+1)^n+\cdots+(x+h)^h=(x+h+1)^n \tag{6.1}$$

并具体地证明了在 $1\leqslant n\leqslant 33$ 时，只有正整数解：

$$\left.\begin{array}{l}1)\text{ 当 } n=1, h=1 \text{ 时}, x=1\\ 2)\text{ 当 } n=2, h=1 \text{ 时}, x=3\\ 3)\text{ 当 } n=3, h=2 \text{ 时}, x=3\end{array}\right\} \tag{6.2}$$

同时还证明了方程（6.1），当 n 为奇数时，除具有 1）与 3）两个正整数解以外，无其他正整数解。根据上述推理，柯召和孙琦猜想：方程（6.1）除解（6.2）以外，无其它正整数解。现已证明，当 $1\leqslant n\leqslant 400$ 时，此猜想成立。对于 $n>400$ 的情形至今尚未见到证明。

还比如，函数

$$f(x)=x+\sum_{n=2}^{\infty}a_nX^n,$$

其中，x 为复变数。该函数在定义域单位圆内（即 $|x|<1$），单值连续，且当 $x=0$ 时，有 $f(0)=0, f'(0)=1$。关于该函数展开式系数 a_2，a_3，…，a_n 的性质问题，自 1909 年以来，吸引了许多数学家的注意力。1916 年，比巴霸赫（Bieberbach）证明了

该函数的系数 $|a_2| \leqslant 2$。由此他提出，$|a_n| \leqslant n(n = 2, 3, \cdots)$ 成立，其中等号当且仅当柯比函数

$$K(x) = \frac{x}{(1 - x)^2}$$

及其旋转时成立。这就是函数论研究中的著名的“比巴霸赫猜想”。

（二）类比法

用类比法提出数学猜想，在数学史上也是常有的事。比如，我们知道，在平面几何中，一个三角形的任意二边之和必大于第三边。后来，人们在反复计算试验的基础上，并在上述事实的启发下，类似地提出了如下猜想：如果 x，y 为大于 1 的自然数，则

$$\pi(x) + \pi(y) \geqslant \pi(x + y),$$

其中 $\pi(x), \pi(y), \pi(x + y)$ 分别表示不超过 $x, y, (x + y)$ 的素数 的个数，这个猜想正确与否，至今没有结论。

又比如，魏尔猜想就是将方程组与方程进行类比后提出来的。事实上，我们知道方程

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots\cdots + a_n = 0$$

有 n 个根，而这 n 个根以及方程的系数 a_1，a_2，a_3，$\cdots$，a_n 的取值范围为全体实数或复数，其数目又是无限多个。假若我们限制方程的系数和根的值，只能在有限个数中选取，那么这种方程是否有解呢？这就是有限域上多变量多项式解的问题。1949 年，魏尔首先证明了这种方程解的个数满足某些条件，然后他利用类比法提出了这些条件对于解方程组也是必须的猜想。这个猜想，于 1974 年被德利涅证明是正确的。

魏尔是在代数方程与代数方程组之间进行类比之后提出数学猜想。那么，在代数方程与非代数方程之间能否也进行类比并提

出数学猜想呢？回答是肯定的。欧拉猜想的提出就是典型的一例。

17 世纪末 18 世纪初，瑞士著名数学家雅克·伯努利（J.Bernoulli，1654—1705）研究并计算出许多无穷级数之和，但对自然数平方的倒数这个无穷级数：

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}+\cdots\cdots$$

的和，一直未能求出。对此他十分关注，曾表示：假如有人能够求出这个他直到现在还未能求出的和并把结果通知给他，他将会十分感激。这个问题，引起了瑞士的另一位著名数学家欧拉（L.Euler，1707—1783）的兴趣。欧拉仔细地研究了这个无穷级数，并发现种种表达式，还计算出 7 位有效数字的和（即 1.644934），但由于这仅是近似值而感到不满意。怎么办呢？他巧妙地采用类比法提出猜想，然后加以严格证明。其具体做法是：首先设 $2n$ 次方程

$$a_0-a_1x^2+a_2x^4-\cdots+(-1)^na_nx^{2n}=0 \tag{6.3}$$

有 $2n$ 个不同的根

$$\alpha_1,\quad -\alpha_1,\quad \alpha_2,\quad -\alpha_2,\quad \cdots,\quad \alpha_n,\quad -\alpha_n.$$

根据多项式因式分解法，得

$$\begin{aligned}&a_0-a_1x^2+a_2x^4-\cdots+(-1)^na_nx^{2n}\\&=a_0\left(1-\frac{x^2}{\alpha_1^2}\right)\left(1-\frac{x^2}{\alpha_2^2}\right)\cdots\left(1-\frac{x^2}{\alpha_n^2}\right),\end{aligned}$$

且

$$a_1=a_0\left(\frac{1}{\alpha_1^2}+\frac{1}{\alpha_2^2}+\cdots+\frac{1}{\alpha_n^2}\right)。 \tag{6.4}$$

其次，考虑下列方程

$$\sin x=0, \tag{6.5}$$

即

$$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots=0。$$

欧拉认为，（6.5）左边有无穷多项，又是无穷次的，所以必有无穷多个根

$$0,\ \pi,\ -\pi,\ 2\pi,\ -2\pi,\ 3\pi,\ -3\pi,\ \cdots\cdots$$

抛弃0这个根，并用 x（即对应于0根的线性因子）除（5.5）的两边，得

$$1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\cdots=0。\qquad(6.6)$$

此方程的根为

$$\pi,\ -\pi,\ 2\pi,\ -2\pi,\ 3\pi,\ -3\pi,\ \cdots$$

这时欧拉大胆地将（6.6）与（6.3）进行类比，从而推断出

$$\frac{\sin x}{x}=1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\cdots$$

$$=\left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{4\pi^2}\right)\left(1-\frac{x^2}{9\pi^2}\right)\cdots$$

又根据（6.4）有

$$\frac{1}{3!}=\frac{1}{\pi^2}+\frac{1}{4\pi^2}+\frac{1}{9\pi^2}+\cdots$$

等式两边乘以 π^2 得

$$\frac{\pi^2}{6}=1+\frac{1}{4}+\frac{1}{9}+\cdots$$

即

$$\frac{\pi^2}{6}=1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots$$

这就是欧拉采用类比的方法推断出雅克·伯努利未能求得的自然数平方的倒数这个无穷级数和。这是一个绝妙的猜想！欧拉自己也表示，这种类比方法是新的且从来没有这样运用过。他又谨慎地作了检查和计算，发现 $\frac{\pi^2}{6}$ 的取小数点后六位数字的近似值与原来的近似值相同。当时，数学界的许多数学家对欧拉的这个推断提出异议。最后，欧拉经过仔细地检验，并成功地证明了用

$\frac{\pi^2}{6}$作为自然数平方倒数这个无穷级数的和是正确的。①

（三）变换条件法

提出数学猜想还有一种常用的方法即变换条件法。所谓变换条件法，就是通过改变某一数学定理的前提条件来提出数学猜想。古希腊数学家欧几里得（Euclid，约前330—约前275）提出并证明了“素数有无穷多”这一著名的数学定理。后来，人们通过变换这一定理条件的方法提出了种种猜想。“孪生素数猜想”就是其中的一个。若 p 是素数，$p+2$ 亦是素数，则称 $(p, p+2)$ 是一对孪生素数。如，（3，5）；（5，7）；（11，13）；（17，19）；（101，103）；（10016957，10016959）；$(10^9+7, 10^9+9)$ 等等，都是孪生素数。孪生素数显然是素数中的一部份，但人们却改变“素数有无穷多”这一定理的条件，提出“孪生素数有无穷多”这一数学猜想。这个猜想至今尚未解决。人们目前已知道的孪生素数是相当多的，大体情况是：在小于 10^5 的自然数中有1224对；在小于 10^6 中有8164对；在小于 3.3×10^7 中有152892对。最大的孪生素数对为$(10^{12}+9649, 10^{12}+9651)$，不仅如此，人们采用同样的方法，又相继提出所谓“三生素数猜想”等。

（四）物理模拟法

有些数学猜想是通过物理模拟并在物理模拟的启示下提出来的。比如，在场站设置的实际问题中，人们归结出这样一个数学问题：对平面上已知 n 个点，把这 n 个点连结起来，如何连线

① 参见G·波利亚：《数学与猜想》（第一卷），科学出版社，1984年版，第17—21页。

才能使总长度最短？为了解决这个问题，人们曾利用物理模拟的方法予以探讨。先选定一块大小适当的细铁丝网，并在给定的几个点的位置上各插一大头针，然后把它放在肥皂水里，最后再轻轻地将铁丝网取出。这时，如果从垂直于铁丝网的方向看去，便可以清楚地看出铁丝网上形成一些网状线，而且从具体测定发现这些线与线之间的结点角① 不小于 120°。过去有人把这个实验称之为“皂膜实验”。在这个实验的启示下，人们提出“在一个平面上 n 点连线总长度最短时其连线间的结点角皆不小于 120°”的猜想。当 $n=3$ 时，我们可以用初等几何的方法证明此推断中的条件不但是必要的，而且也是充分的。但对于 $n>3$ 时，其条件仅仅是必要的，至于充分条件至今尚未找到。

（五）联系观察法

在数学研究中，有时出现这样一种情况，即通过联系观察发现某种带有规律性的现象，然后再从理论上加以探讨，究其真伪性。这种通过联系观察而发现的某种带有规律性的东西，亦可称之为数学猜想。这里，仅以波利亚《数学与猜想》一书中所讲到的事实为例，根据我们的理解和需要来作具体的分析和说明。

今考虑用位置为一般性的点、直线、平面分别分割直线、平面、空间，并力图发现其分割部分数的规律。首先，我们来看几个不同的点，能把直线分成多少部分？很显然，1 个点将直线分为 2 个部分；2 个点将直线分为 3 个部分；3 个点将直线分为 4 个部分；一般地，n 个点将直线分为 $n+1$ 个部分。其次，我们再来看几条直线将平面最多能分割成多少部份？通过实际观察不难看出，1 条直线将平面分成 2 部分；2 条直线将平面分成 4 部分，3 条直线将平面分成 7 部分；4 条直线将平面分成 11 部分

① 从某一点出发的射线间的夹角称为关于这一点的结点角。

等。这里有何规律呢？几条直线将平面分为多少部分呢？一下子很难看出。为发现规律，我们可以将上面观察出来的结果，作如下处理：

直线条数	平面被分割的份数
1	2 = 1 + 1 + 0
2	4 = 1 + 2 + 1
3	7 = 1 + 3 + 3
4	11 = 1 + 4 + 6
⋮	⋮
n	$p_n = 1 + n + a_n$

从此表右边的等式中，仔细观察便可发现，每个等式右端第一项均为 1，第二项均与直线的条数相同，就是说 n 条直线即为 n。如果第三项 a_n 亦能用 n 表示出来，那么这个分割份数的一般公式即可写出来了。我们再来仔细观察可发现第三项各数值中，相继后项减前项所得数值是一个自然数列。事实上，设 $a_1 = 0$，$a_2 = 1$，$a_3 = 3$，$a_4 = 6$，……，则 $a_2 - a_1 = 1$，$a_3 - a_2 = 2$，$a_4 - a_3 = 3$，……。这时我们可以猜想，一般地 $a_n - a_{n-1} = n - 1$，现将等式左右两端分别相加得

$$a_n - a_1 = 1 + 2 + 3 + \cdots\cdots + (n - 1)。$$

又因为 $a_1 = 0,\ 1 + 2 + 3 + \cdots + (n - 1) = \dfrac{n(n-1)}{2}$，

所以

$$C_n = \frac{n(n-1)}{2}。$$

这样一来，即可得到一般性推断：n 条直线可将平面分割的份数为

$$p_n = 1 + n + \frac{n(n-1)}{2}。$$

这只是一个猜想，是否真正成立尚需进行证明。用数学归纳法不难证明这个猜想是正确的。

n	Q_n	p_n	l_n
0	1	1	1
1	2	2	2
2	4	4	3
3	8	7	4
4	15	11	5
⋮	⋮	⋮	⋮

最后，我们来讨论 n 个平面将空间最多能分割多少部分这个问题。如同前面一样，亦可先做简单的具体分割。1 个平面将空间分为 2 部分；2 个平面将空间分为 4 部分；3 个平面将空间分为 8 部分；4 个平面将空间分为 15 部分（由四个面围成的四面体，其内有限部分为 1 部分；与四面体共面的无限部分有 4 部分；与四面体共一边的无限部分有 6 部分；与四面体共一顶点的无限部分有 4 部分。总共为 15 部分）等。看 1，2，4，8，15，…有何规律？不易看出。为此，我们可将所观察到的点分割直线，直线分割平面与平面分割空间的具体数值联系起来考虑，并列入左表，再看一看相互之间有什么规律可循。其中，n 表示分割元素的个数；l_n 表示直线被 n 个点分割的份数；p_n 表示平面被 n 条直线分割的份数；Q_n 表示空间被 n 个平面分割的份数。

我们观察一下右三列中相邻两列数字的关系。可以发现，从第二行开始，每个数值均等于本列中上一行的数值与其同行右侧数值之和。比如，

$$\left.\begin{matrix}4,\ 3\\7\end{matrix}\right\}7=4+3;\qquad \left.\begin{matrix}4,\ 4\\8\end{matrix}\right\}8=4+4;$$

$$\left.\begin{matrix}7,\ 4\\11\end{matrix}\right\}11=7+4;\qquad \left.\begin{matrix}8,\ 7\\15\end{matrix}\right\}15=8+7;$$

依据上述观察结果，我们可以提出下列猜想：n 条直线分割平面

的份数等于 $n-1$ 条直线分割平面的份数加上 $n-1$ 个点分割直线的份数；n 个平面分割空间的份数等于 $n-1$ 个平面分割空间的份数加上 $n-1$ 条直线分割平面的份数。如果这个猜想是对的，那么我们就可以通过这个表算出 n 条直线分割平面的份数和 n 个平面分割空间的份数。诸如，当 $n=5$ 时，平面被直线分割的份数为 $p_5=11+5=16$，空间被平面分割的份数为 $Q_5=15+11=26$；当 $n=6$ 时，$p_6=16+6=22$，$Q_6=26+16=42$；当 $n=7$ 时，$p_7=22+7=29$，$Q_7=42+22=64$ 等等，详见下表：

n	Q_n	p_n	l_n
⋮	⋮	⋮	⋮
5	26	16	6
6	42	22	7
7	64	29	8
8	93	37	9
9	130	46	10
10	176	56	11
⋮	⋮	⋮	⋮

联系观察法在数学猜想的提出中，起着极其重要的作用。这不仅是因为它本身有发现猜想的功能，而且也是由于各种提出数学猜想的方法中大都离不开它。正如著名数学家欧拉在谈到数论研究的情况时所指出的那样："今天人们所知道的数的性质，几乎都是由观察发现的，并且早在用严格论证确认其真实性之前就被发现了。甚至到现在还有许多关于数的性质是我们所熟悉而不能证明的；只有观察才使我们知道这些性质。因此我们认识到，在仍然是很不完善的数论中，还得把最大的希望寄托于观察之

中；这些观察将导致我们继续获得以后尽力予以证明的新的性质”。[①] 事实上，数论中的大部分猜想是靠联系观察法（当然包括归纳法）提出的，而其他数学分支中数学猜想的提出，也常常与这种方法有密切关系。

（六）逐级猜想法

在数学中，有些猜想长期不能肯定其正确与否。对这些猜想，有时先假定它是正确的，然后在这个假定的基础上再提出新的推断，从而由已有的猜想又得到了新的猜想。比如，前面已经讲过，“孪生素数有无穷多个”是一个猜想，“三生素数有无穷多个”也是一个猜想。这两个猜想至今尚未解决，但后来人们在假定这两个猜想是正确的前提下，又运用“归纳法”，进而提出“n 生素数猜想”。所谓“n 生素数猜想”，就是假定 $L_1 < L_2 < \cdots\cdots < L_{n-1}(n > 1)$ 是 $n-1$ 个自然数，P 是素数，且 $P+L_1, P+L_2, \cdots\cdots, P+L_{n-1}$ 也都是素数，则称 $(P, P+L_1, \cdots, P+L_{n-1})$ 是一个 n 生素数组，如果对于任意素数 q，n 个整数 0，L_1，…，L_{n-1} 对模 q 互不同余的个数都小于 q，那么 n 生素数组有无穷多个。

还比如，从连续统假设到它的推论，也是一个从猜想到猜想的例证。我们知道，由全体自然数构成的集合具有“可数”的势，而由全体实数构成的集合具有“连续统”的势。“连续统”的势。大于“可数”的势。那么，“可数”的势与“连续统”的势之间，是否还有其他的势呢？集合论的奠基者康托猜想：在“可数”的势与“连续统”的之间没有其他的势，这就是著名的“连续统假设”。后来人们在假定这个假设是正确的条件下，又运用“演绎法”（引号的意思与上例中“归纳法”同）得到 82 个推

---

① 转引自 G·波利亚：《数学与猜想》，科学出版社，1984 年版，第 1 页。

论。可见这些推论是从猜想中得到的新的猜想。

再比如，1921 年，克拉莫（H.Cramer）在假定“黎曼猜想”成立的条件下，证明了：当 $x=p_n$，$y=p_n^{1/2}\log p_n$ 时，在区间〔x，$x+y$〕中必定有素数存在，从而将杰波夫猜想的研究向前推进了一步。这里，克拉莫证明了的命题，显然具有由猜想到猜想的性质。

上面讲到了提出数学猜想的几种主要方法。这些方法在实际运用中，有时是孤立地进行，但也有时几种方法一同使用。比如，前面讲到的观察法的事例中，不仅用了观察法，而且还有不完全归纳法的作用。另外，数学猜想并不是科学家随意提出来的，而是根据数学研究的实际需要，并在数学研究发展到一定程度时产生的。数学猜想作为数学研究的一种方法和发展形式，及时发现和提出它，对数学科学的进步无疑是有重要意义的。

七、闪光的并非都是金子

——判定数学猜想真伪性的几个途径

数学猜想具有两个明显的特点：一是具有一定的科学性，因为它是以某些已知的事实材料与数学知识为依据，有时还经过一定的理论推证，因而它与毫无根据的武断、臆测、胡思乱想有着本质的区别；二是具有某种假定性，因为它没有经过全面而严格地理论证明和实践检验，是一种猜测性的推断。数学猜想是科学性与假定性的辩证统一。既然数学猜想本身具有科学性和假定性，那么究其结果，既可能被肯定，也可能被否定，还可能是不可判定的。那么，如何从理论上进行判定呢？归纳起来有以下几个途径。

（一）举例否定

对于一个可判定的命题，要么证明它是正确的，肯定它；要么证明它是错误的，否定它。对于某个数学猜想如果能举出一个反例，那么这个数学猜想便被否定了。比如，前面我们讲了，1664 年费尔马研究形如 $F(n)=2^{2^n}+1$ 的数，并根据 $n=1,2,3,4$ 时，$F(n)$ 均为素数，提出猜想：对于任何自然数 n，形如 $F(n)=2^{2^n}+1$ 的数都是素数。但是，事隔 68 年，即 1732 年，欧

拉举出一个反例：当 $n=5$ 时，$F(5)=2^{2^5}+1=4294967297$，这个数不是素数，而是合数，因 $4294967297=6700417\times641$，即可被 641 整除。这样，费尔马猜想就被否定了。

还比如，欧拉方阵猜想（亦称欧拉方阵问题），也是通过举反例被否定的。1782 年，荷兰的一个杂志发表了一篇关于魔方阵的文章，文中提到：传说在 18 世纪，普鲁士的腓特烈大帝举行一次阅兵式，计划挑选 36 名军官组成一个方队作先导队。普鲁士当时有六个部队，腓特烈大帝要求，从每个部队中选派出 6 个不同级别的军官各一名，并使这 36 名军官排成的方队中，必须每一行每一列都有各部队各级别的代表。阅兵司令按照腓特烈大帝的命令去排方阵，但直到阅兵时也未能排出。对此，腓特烈大帝大怒，并撤了阅兵司令的职务。后来，腓特烈大帝亲自来排也未获成功。瑞士数学家欧拉对这个“6 阶方阵问题”十分感兴趣。当时，他已是 75 岁高龄的老人了。他一生中解决了不知多少令人望而生畏的数学难题，可是万万没想到竟被这个看来很简单的问题给难住了。开始，他从简单情况入手，把 3 阶、4 阶、5 阶的情况按要求把方阵排出来了，如下所示。

$$\begin{array}{ccc} A_a & B_c & C_b \\ B_b & C_a & A_c \\ C_c & A_b & B_a \end{array} \qquad \begin{array}{cccc} A_a & B_b & C_c & D_d \\ D_b & C_a & B_d & A_c \\ B_c & A_d & D_a & C_b \\ C_d & D_c & A_b & B_a \end{array} \qquad \begin{array}{ccccc} A_a & B_b & C_c & D_d & E_e \\ E_d & A_e & B_a & C_b & D_c \\ D_b & E_c & A_d & B_e & C_a \\ C_e & D_a & E_b & A_c & B_d \\ B_c & C_d & D_e & E_a & A_b \end{array}$$

然而，对 6 阶这种方阵，他绞尽脑汁也没有排出来，这时，欧拉想：6 阶这种方阵是否不存在？他又仔细研究发现，2 和 6 都是被 2 整除而不能被 4 整除的数（欧拉称这种数为半偶数），其一般形式为

$$n=4m+2(m=0,1,2)$$

于是，欧拉猜想：半偶数的方阵是不存在的。第二年，即 1783

年，欧拉去逝了。后来人们就把这类半偶数方阵称为“欧拉方阵”。自欧拉提出这一猜想后，一个多世纪中，不少数学家为证明它而耗费了极大的精力，付出了艰苦的劳动，并经历了多次反复。1901 年，丹麦数学家彼得尔森用几何方法证明了 $n=6$ 时欧拉方阵猜想是正确的；但是 1905 年，法国数学家塔里却指出他的证明是错误的；1910 年，德国数学家维尔尼克用代数方法证明了欧拉方阵猜想是正确的，但是，美国数学家麦克尼许又指出他的证明有错误；1922 年，麦克尼许用拓扑的方法证明了欧拉方阵猜想是正确的，但是，1942 年，德国数学家勒维发现麦克尼许的证明同样是不可靠的。证明欧拉方阵猜想正确的努力长期遭到失败之后，数学家们开始从相反方向去考虑问题，即力图证明欧拉方阵猜想的不正确性。1959 年 4 月，印度数学家玻色和史里克汉德证明当 $n=22$ 时方阵是存在的。这个反例便推翻了上述欧拉方阵猜想。这个事件引起了数学界的震动！不久，玻色和史里克汉德又证明了除 $n=2$，6，14，26 以外，对 $n\geqslant 3$ 的任意 n 都存在方阵。就在这篇论文交付印刷时，美国数学家派克又证得了 $n=14$，26 的方阵是存在的。这样一来，只剩下 $n=2$，6 时方阵不存在了。从而最后明确了，36 名军官的方阵确实是不存在的。

再比如，19 世纪末，泰特提出关于“任何 3-连通的三次平面图是哈密尔顿的”断言，也就是著名的泰特猜想。这个猜想的大意是，如果把由点和线组成的图能够画在平面上，且线与线之间除了有公共的端点外没有任何交点，这样的图在图论中叫做“平面图”。如果把图中的每个点视作一个城市，那么联结两个点的线便可看作是交通线。1859 年，哈密尔顿提出了相当于下面这样一个问题：能不能找到一条旅行路线，从一个城市出发，沿着交通线经过每个城市恰好一次，再回到原出发地？如果能找到这样一条旅行路线，我们就称这样的图为一个哈密尔顿图。并不是每个平面都是哈密尔顿图，但许多具有 3-连通的三次平面图

是哈密尔顿的。这是不是一个普通规律呢？1946 年，托特给出了一个 46 个点的具有上述性质的平面图的反例，从而证明了泰特猜想是不对的。

（二）逐次趋近

数学猜想中有不少世界著名的难题，对于这些世界难题，人们常常设法先证明它的一种减弱的命题，然后一步一步地向它逐次趋近。数学发展史上有许多这样的例子。

例 1．前面提到的哥德巴赫猜想，自 1742 年被提出后，广大数学家陆续作出了越来越接近最后解决（假定以偶数 = （1 + 1）来表示）的成果：

1924 年，拉德马哈尔证明了：偶数 = (7 + 7)；

1932 年，爱斯斯尔曼证明了：偶数 = (6 + 6)；

1938 年，布赫斯塔勃证明了：偶数 = (5 + 5)；

1940 年，布赫斯塔勃证明了：偶数 = (4 + 4)；

1957 年，维诺格拉朵夫证明了：偶数 = (3 + 3)；

1957 年，王元证明了：偶数 = (2 + 3)；

1962 年，潘承洞证明了：偶数 = (1 + 5)；

1962 年，王元和潘承洞证明了：偶数 = (1 + 4)；

1965 年，布赫斯塔勃等证明了：偶数 = (1 + 3)；

1973 年，陈景润证明了：偶数 = (1 + 2)。

例 2．1859 年，德国数学家黎曼（B.Riemann，1826—1866），一连提出六个猜想。第一、第三、第四个猜想已于 1892 年由法国数学家哈达马证明；第二、第六个猜想已于 1894 年由曼高尔特解决。后来只剩下第五个猜想，即函数

$$\xi(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

（其中 $s = \sigma + ti$ 为复数）的零点全部落在复平面的一条直线 $\sigma =$

$\frac{1}{2}$上。一百多年来，许多数学家就其减弱的命题进行了试证，并取得了一系列重要成果，逐步向最后结果靠近。为了说明这种情况，我们先来讨论一下黎曼 $\zeta(s)$ 函数在一个有限范围内的零点个数问题。设这个范围是 s 的虚数部分的绝对值不超过 T，实数部分在 0 与 1 之间。用 $N(T)$ 表示 $\zeta(s)$ 在这一范围内的零点个数。若 s 的范围改为虚数部分的绝对值不超过 T，实数部分等于 $\frac{1}{2}$，用 $N_0(T)$ 表示 $\zeta(s)$ 在其中的零点个数。假如我们能证明 $N(T) = N_0(T)$，那么黎曼猜想就被证明了。

下面我们来介绍一下目前研究的进展。1948 年，塞尔伯格（Selberg）证明了一定存在这样一个常数 c，使 $N_0(T) \geqslant cN(T)$ 成立。对其中 c 取什么样数值他没有明确，实际上按他的方法只能求出一个十分小的 c。许多数学家力图找到一个较大的 c。如果 c 可取为 1，则黎曼猜想便被证明是正确的。这是因为 $N_0(T) \leqslant N(T)$ 是显然的，如果有 $N_0(T) \geqslant N(T)$，那么 $N_0(T) = N(T)$ 就是必然的了。1973 年，莱文生（Levinson）证明了：$N_0(T) > N(T)/3$。1974 年，他又改进为 $N_0(T) > 0.3474N(T)$。1980 年，我国数学家楼世拓，又改进了上述结果。

由于“$\xi(s)$ 的非平凡零点[①] 全部在 s 的实数部分等于$\frac{1}{2}$的这条直线上”与“在上述这条直线之外 $\zeta(s)$ 没有非平凡零点”这两个命题是等价的，所以我们可以证明，在 $s = \sigma + ti$ 的实部 σ 大于或等于 1 时，$\xi(s)$ 没有零点，但却不能证明：对于任何一个比 1 小的数 a 来说，σ 大于或等于 a 时，$\zeta(s)$ 没有零点。我们只好考虑再把命题减弱一些，即希望在上述范围内零点个数

---

① 当 S = -2，-4，…时，$\xi(s) = 0$，称这些零点为“平凡零点”，此外的零点均称“非平凡零点”。

不很多。这就是所谓零点密度问题，对于一个大于$\frac{1}{2}$的 a 来说，在 s 的实数部分σ 满足$1>\sigma\geqslant a$，s 的虚数部分t 满足$|t|\leqslant T$ 的这个范围之中，$\zeta(s)$的零点的个数记为$N(a,T)$。对于$N(a,T)$的估计就称为零点密度估计。目前这方面的最好结果是赫克斯雷（Huxley）在 1923 年得到的结果：$N(a,T)\leqslant T^{2.4(1-\sigma)}$。①从上述的一系列研究成果看，确实逐渐向黎曼猜想的最终解决趋近，但仍相距甚远。

例 3. 1932 年，勒默（Lehmer）提出猜想：不存在合数 n 使得$\varphi(n)\mid n-1$，即推断数论函数方程 $k\rho(n)=n-1(k\geqslant 2)$ 无正整数解。其中，$\varphi(n)$是欧拉函数，表示不超过 n 且与 n 互素的正整数的个数。1932 年，勒默本人证明了：这样的 n 如果存在，则 n 至少是 7 个不同的素数的乘积；1962 年，柯召与孙琦证明了 n 至少是 12 个不同素数的乘积；1963 年，柯召与孙琦证明了 n 至少是 13 个不同素数的乘积；1970 年，有的数学家证明了 n 至少是 11 个不同素数的乘积；1975 年，克萧（Kishore）又证明了 n 至少是 13 个不同素数的乘积。证明这个猜想的难度很大，现有不少的数学家正在积极研讨着它。

例 4. 1950 年，欧德斯（Erdös）提出猜想：对于一切 $n>1$ 的正整数，方程

$$\frac{4}{n}=\frac{1}{x}+\frac{1}{y}+\frac{1}{z} \qquad (*)$$

皆有正整数解 x，y，z。史劳斯（Strauss）进一步猜想：当 $n\geqslant 2$ 时，方程$(*)$的解 x，y，z 满足$x\neq y$，$y\neq z$，$x\neq z$，并且对 $2<n<5000$ 证明了这个论断是正确的。1963 年，柯召、孙琦与张先觉证明了欧德斯猜想与史劳斯猜想是等价的，且证明了 $n<400000$ 时，这个猜想成立。后来，杨曼托（Yamanoto）又

① 参见楼世拓：《关于黎曼猜想》，《自然杂志》，1980 年 5 期，第 355—357 页。

把这个结果发展到 $n<10^7$。

例 5，采取逐次趋近的方法证明数学猜想也并不都是由特殊的命题一步一步地、机械地进行，有时是“一般”与“特殊”交错进行。比如，证明“彭加勒猜想”（在 n 维空间中的一个点集，若是 $n-1$ 连通的紧致流形，则必定是 n 维球）的情形就是如此。关于这个猜想，当 $n=1$，$n=2$ 时，人们早已知道其结果是正确的。1960 年，美国数学家斯梅尔（S. Smale）证明了当 $n\geqslant 5$ 的一般情形也是对的。但是当 $n=3$，$n=4$ 的特殊情形，却长期得不到解决。一直到 1981 年，美国数学家弗里德曼（M. Freedman）才证明了当 $n=4$ 时，彭加勒猜想是对的。又过 5 年，到 1986 年，葡萄牙数学家莱戈（E. Rêgo）和英国数学家罗克（C. Rourke）最后证明了当 $n=3$ 时，彭加勒猜想是成立的。这样，自 1904 年彭加勒提出这一猜想，到 1986 年最后解决这一猜想，共经历了 82 年的时间，人们是通过“特殊——一般—特殊”的曲折途径获得成功的。

（三）命题转化

有些数学猜想，采用直接证明的方法长期得不到进展，怎么办呢？人们往往选取命题转化的途径。其具体作法有二种。第一是转化为等价命题，即要证明某个数学猜想，先提出与其等价的命题，然后证明这个等价命题，从而原数学猜想得证。比如，1882 年，德国数学家康托尔提出“连续统假设”。1934 年，为了证明这一著名的数学猜想，谢宾斯（Sierpinshi）曾列举了 12 个与它等价的命题，还列出了关于连续统假设的 82 个推论。后来人们在此启发下，把“连续统假设”的研究向前推进了一大步。又比如，关于牛曼（Newman）猜想也是这样。此猜想的内容是整数 $m+1$，$m+2$，…，$m+n$ 总可以重新排列成 $m+i_1$，$m+$

i_2，…，$m+i_n$，使 $(m+i_j, j)=1$，①$i=1$，2，…，n。对这个数学猜想，1963 年戴金（Dayhin）和贝弥斯（Baines）证明了在 $n=m$ 时此猜想成立。同年，柯召、孙琦、尹文霖证明了在 $n\leqslant 17016$ 时此猜想成立。1971 年，卡福特（Chvétai）证明了 $n\leqslant 1002$ 时此猜想成立。但至今未能得到最后证明。我国数学家闵嗣鹤，为了证明这一猜想，曾证明了与它等价的命题。这个等价命题是：若 $b\geqslant 0$，而 $0<a_1<a_2<\cdots\cdots<a_h\leqslant n$，$b<b_1<\cdots\cdots<b_k\leqslant b+n$ 都是整数，且 $(a_i, b_j)>1$，则 $k+h\leqslant n$。对于不太大的 n，可以用这个命题来检验牛曼猜想是否正确。闵嗣鹤曾用这个结果检验了 $n=10$ 和 $n=100$ 时，此猜想是成立的。

运用命题转化来证明数学猜想的第二个具体作法是，要证明某个数学猜想，先证明另一个数学猜想成立，然后用这个数学猜想推证出原数学猜想的成立。比如，1916 年，比巴霸赫提出著名的“比巴霸赫猜想”。六十多年来，许多数学家作出了一系列成果，但都未能最后解决。一直到 1984 年，被路易斯·德布朗格斯（Louis de Branges）解决了，引起了数学界的震惊！他是怎样解决的呢？他是先证明了另一个米林（Milin）于 1971 年提出的更强的猜想，因为这个猜想隐含了比巴霸赫猜想，故通过这个间接的途径证明了比巴霸赫猜想。

（四）反　证　法

以上讨论的证明方法，基本上属于非逻辑方法。其实，在数学猜想的论证中，也经常采取逻辑论证的方法。像反证法就是其中的一个。什么叫反证法呢？所谓反证法就是用否定待证结论而导出矛盾来肯定待证结论的一种推理方法。这种推理方法的基本思想是，在肯定前提的情况下，从否定待证结论出发，根据已知

① $(m+i_j, i)=1$ 表示整数 $m+i_j$ 与 j 的最大公约数是 1，即它们互素。

事实（前提或公理、定理等），进行正确的逻辑推理，最后导出与已知事实矛盾，而已知事实是正确无疑的，因此这种矛盾说明了“否定待证结论”是不正确的，从而从反面肯定了待证结论是正确的。比如，早在公元前人们就提出“素数有无穷多”这一猜想。古希腊数学家欧几里得用反证法证明了这个问题。事实上，假定素数只有有限个，如有 k 个：P_1，P_2，…，P_k。这样，$P_1P_2\cdots P_k+1$ 一定可分解为某些素数的乘积，如分解为 $P_{e1}P_{e2}\cdots P_{ei}$，就是说 $P_1P_2\cdots P_k+1$ 可被 P_{e1}，$P_{e2}\cdots P_{ei}$ 整除，即 $M=\frac{P_1P_2\cdots P_k+1}{P_{e1}P_{e2}\cdots P_{ei}}$ 是整数。又由于 $P_1P_2\cdots P_k$ 是所求素数之积，故 $P_1P_2\cdots P_k$ 定被 $P_{e1}P_{e2}\cdots P_{ei}$ 整除。但是，

$$\frac{P_1P_2\cdots P_k+1}{P_{e1}P_{e2}\cdots P_{ei}}=\frac{P_1P_2\cdots P_k}{P_{e1}P_{e2}\cdots P_{ei}}+\frac{1}{P_{e1}P_{e2}\cdots P_{ei}}。$$

因为等式右端第一项是整数，第二项不是整数，故 M 不是整数。这就与前面说 M 是整数矛盾，从而证明了原假定是错误的，亦即证明了“素数有无穷多”这一猜想是正确的。

又比如，“欧氏第五公设是可证的”这一数学猜想也是巧妙地运用反证法解决的。下面，我们就来分析一下这一历史过程。

早在公元前3世纪，古希腊数学家欧几里得等，在系统地总结了古代人们在长期实践中积累起来的数学知识的基础上，完成了数学巨著《几何原本》。《几何原本》把公认的一些事实列为二十三个定义、五个公设和五个公理，并以此为前提，用演绎推理的方法证得了一系列定理，从而形成了世界上最早的一部公理化数学著作。它是数学发展史上的一项重大成果。然而，在开始以及以后相当长的时间里，人们对其中作为推理前提的“第五公设”① 是有怀疑的。因为公设或公理一般说来是通过实践反复证

① 欧氏第五公设的内容是：若两条直线与第三条直线相交，且在同一个侧所构成的两个同侧内角之和小于二直角，则把这两条直线向这一侧适当延长之后一定相交。

明了的公认事实，具有“自明性”，所以就表现为在数学上不加证明也无法证明的东西。而“第五公设”看起来比较复杂，不像其他公设、公理那样便于在实践中直接检验和不证自明，加之《几何原本》中的前二十八个命题的证明均未用到它。这些都促使人们考虑：“第五公设”是否不是“公设”而是可以给出数学证明的“定理”？于是，人们就提出了“欧氏第五公设可证?”这样一个数学猜想，并有许许多多的数学家参与了试证工作。但是，结果都失败了。一直到19世纪初，人们从屡屡的失败中渐渐得到启示，进而考虑：欧氏第五公设是否不可证？如果从理论上证明欧氏第五公设不可证，那么数学家为之奋斗二千多年的试证工作不就得到结果了吗？这样，人们便开始从原来试证“可证”转向“不可证”这一命题。或者说，为了解决“欧氏第五公设可证”这一猜想又进而提出和试证“欧氏第五公设不可证”这样一个新的猜想。

19世纪20年代，德国的高斯(C.F.Gauss, 1777—1855)，俄国的罗巴切夫斯基（Н.И.Лобачевский，1792—1856）和匈牙利的亚·鲍耶（J.Bolyai，1802—1860）这三位数学家，在不同的国度里，几乎同时对这一新的“不可证”猜想进行了深入的研究，并以各自独特的方式，巧妙地运用反证法证明了它。这里，仅就罗巴切夫斯基证明此猜想的基本思想来说明问题的实质。

1826年2月23日，罗巴切夫斯基在喀山大学物理数学系的学术会议上宣读了他关于平行线理论的论文《几何学原理扼要阐释及平行线定理的严格证明》，证明了“欧氏第五公设不可证”这一猜想是正确的。正如他在1835年发表的《新几何原本》中所说的那样：“大家知道，直到今天为止，几何学中的平行线理论还是不完全的。从欧几里得时代起，两千多年的枉然努力，使我们不得不怀疑，人们要证明的这个真理，还不能由原有概念中推出，类似于物理定律，只有经验，例如天文观测，才能验证它。我的猜测的正确性已经得到证实，我认为最困难的这个问题

完全得到了解决，我在1826年谈到了这个论证。”① 那么，罗巴切夫斯基到底是怎样证明的呢？概括地说，就是用反证法证明的。事实上，他的证法大体分为三步：

第一，以《几何原本》中的定义、公理、公设（不包括第五公设）以及与第五公设无关的定理（如《几何原本》中的前28个命题）为基础，作成一个基本的公理系统，并假定第五公设是这个公理系统的推论；

第二，否定第五公设，并给出其否命题：“过已知直线外一点，至少可以引两条直线与已知直线不相交”，同时将此否命题加入上述基本公理系统中去，作成一个扩大了的公理系统，那么，这个扩大的公理系统中必然会有互相矛盾的命题出现（如第五公设与它的否命题就是其中的一对矛盾）；

第三，证明这个扩大的公理系统中没有矛盾。

这三步恰好体现了反证法。因为在作出基本公理系统后，假定第五公设是这个系统的推论，也就是令“欧氏第五公设可证”，这显然是对“欧氏第五公设不可证”这一待证结论的否定。由此出发又得出在扩大的公理系统中，必然存在矛盾的结论。但是，在这种假定条件下的逻辑推演，“其结果没有得到任何矛盾”。于是，就出现了“必然存在矛盾”但又“得不到矛盾”这样一个“矛盾”。这只能说明原来假定“欧氏第五公设可证”是不正确的，因而“欧氏第五公设不可证”这一猜想是正确的。正因为在扩大的公理系统中推不出矛盾，所以在证明的过程中便形成了一整套系统的逻辑上无矛盾的几何理论，即非欧几何学。

这里，还应强调指出的是，如果说，欧氏几何是以归纳为开端，演绎为手段，构造其理论体系，充分体现了归纳法的创造性功能，那么非欧几何却是同时以演绎法为开端和手段，建立起来

① 转引自A·П·诺尔金：《几何学基础》，国立技术出版社，1956年版，第61页。

的理论体系，深刻反映了演绎法也能开辟数学新天地。这一事实，有力地说明了过去人们对演绎法作用的传统看法是有局限性的。过去有人认为，归纳法能够“促成知识的真正进步”，“富有创造性功能”，而演绎法“只能扩大已有的理论成果”，“不可能导出新的概括”，“不可能在科学上作出较大的进展”，一句话，就是没有开辟新领域的作用。事实证明这种看法是不符合实际的，是片面的。

判定数学猜想的真伪性，除上述四种途径以外，还有机器证明（详见本书第五章）以及运用各种常规方法等。当某猜想长久得不到解决，人们还往往采取先进行能行性分析，如果断定是“不可判定性”的，则另寻新方法，开辟新领域，以求解决之。

八、千淘万漉始到金

——数学猜想的艰难性

在本书的前言中，我们就谈到了数学猜想一般说来是一时解决不了的数学难题，就是说，它具有明显的艰难性。这种艰难性，主要表现在以下三个方面。

（一）有一个逐步完善的过程

数学研究的实际表明，数学猜想与其他任何数学问题的提出一样，并非一帆风顺，往往要经过一个酝酿、萌发、修改和完善的过程。现以凸多面体的面数（F）、顶点数（V）与棱数（E）的关系公式：$F+V=E+2$ 的提出为例予以说明。

我们首先选择一些个别的多面体：（1）三棱锥，（2）立方体，（3）五棱台，（4）八面体，（5）十二面体，（6）二十面体，（7）塔顶体，（8）截角立方体，（9）截角五棱柱。

然后，数清它们各自的面数（F），顶点数（V），和棱数（E），并列出下表（见下页）。

现在，我们来观察一下此表中的各数字，看它们之间有何带有规律性的关系？开始时，不易一下子看出什么，只好试考虑一些可能。比如，V 是否随 F 的增加而增加？仔细观察可看出，表中的(1)、(2)、(3)，依次确有这种趋势，即 F 数为 $4\to6\to7$，

则 V 数为 4→8→10。但再看(4)与(7)，却出现 F 增加了，但 V 并没有增加，故此种可能不成立。又比如，E 是否随 F 或 V 的增大而增加？观察后可发现：从(7)到(8)，V 是从 9 到 10，但 E 却是从 16 到 15，前者增加了，而后者却减少了；从(3)到(4)，F 从 7 到 8，但 E 却是从 15 到 12，前者增加了，而后者却减少了，故这种可能也不成立。那么，到底 F，V，E 这三者的关系有何规律呢？再仔细观察，就可以发现，E 总是随 $F+V$ 的增加而增加的，并通过计算可知，这个增加的具体公式是：$F+V=E+2$。于是，我们可以运用归纳法提出猜想：一切多面体的面数(F)、顶点数(V)与棱数(E)的关系均为：$F+V=E+2$。

	多面体	面数（F）	顶点数（V）	棱数（E）
(1)	三棱锥	4	4	6
(2)	立方体	6	8	12
(3)	五棱台	7	10	15
(4)	八面体	8	6	12
(5)	十二面体	12	20	30
(6)	二十面体	20	12	30
(7)	“塔顶”体	9	9	16
(8)	截角立方体	7	10	15
(9)	截角五棱柱	8	12	18

一个猜想提出后，人们总是要考虑：这一猜想是否理想？就是说，是否存在一些很容易就发现的相反的情形？如果存在，就应对原提出的猜想进行修改，使之更加完善。为此，我们再来进行一些检验。

我们依据这个想法，可找到一个镶嵌画的框架状多面体。取

一条三棱形的杆，把它分割成四段，把它装配成一个框架状多面体，并假设这个框架是放置在平板上的。先计算棱数，水平棱数有 $4\times3=12$，不水平的棱数也有 $4\times3=12$，所以总数 $E=12+12=24$。再计算面和顶点数，我们发现 $F=4\times3=12$，而且 $V=4\times3=12$。结果，$F+V=24$ 与 $E+2=26$ 不相等，即对这个多面来说，上述猜想并不成立。前面讨论的多面体与这种多面体有什么不同呢？根本区别在于前者是凸多面体，而后者是非凸多面体。因此，我们只好把原来的猜想加以限制，改为：一切凸多面体的面数（F）、顶点数(V)与棱数(E)的关系为 $F+V=E+2$。这一猜想实际上是著名数学家欧拉提出的，所以有时被人们称为“欧拉公式”。

总之，一个好的数学猜想的提出确是不容易的，它不仅要求提出者具有雄厚的数学基础知识和较高的分析洞察能力，而且要求提出者还要深思熟虑、不厌其烦地检验与修改，使其不断完善。数学史上一些重大的数学猜想，为什么大部分出自大数学家之手，原因也就在这里。

（二）时间长与途径曲折

数学发展史表明，数学猜想尤其是一些重大数学猜想的研讨和解决，是极其困难的，其主要表明有二个方面。

1. 时间长

历史表明，数学猜想的解决，有的需要几年、十几年，也有的需要几十年、几百年的，还有的需要一千年、两千年……。比如，魏尔猜想自 1949 年提出到 1974 年德利涅证明是正确的，经过了 25 年；欧德斯猜想自 1950 年提出至今已经过 36 年尚未解决；勒默猜想自 1932 年提出至今已经过 54 年尚未解决；泰特猜想自 19 世纪末提出到 1946 年托特举出反例否定了它，经过了近

50 年；费尔马猜想自 1664 年提出到 1732 年欧拉举出反例否定了它，经过了 68 年；连续统假设自 1882 年由康托尔提出至今经过 104 年尚未解决；黎曼于 1859 年提出六个猜想，第一、第三、第四个由哈达玛于 1892 年解决，经过了 33 年，第二、第六个由曼高尔特于 1894 年解决经过了 35 年；第五个至今已有 127 年尚未解决；四色猜想自 1840 年由莫比乌斯提出至 1976 年由阿佩尔、黑肯获证，经过了 136 年；凯特兰猜想自 1842 年提出至今已有 144 年尚未解决；哥德巴赫猜想自 1742 年提出至今已有 244 年尚未解决；默森尼猜想自 1644 年提出至今已有 342 年尚未解决；“费尔马大定理”自 1637 年提出至 1994 年由维尔斯证明，经过了 357 年；“欧氏第五公设可证”猜想自古希腊提出后到 19 世纪 20 年代由高斯、罗巴切夫斯基、亚·鲍耶最后解决，经过了 2000 多年等等。

2. 途径曲折

一个数学猜想从提出到最后解决，有时即使由世界各国优秀数学家进行研讨，也往往要经过一个漫长而曲折的途径。比如，著名的比巴霸赫猜想，自 1916 年提出后，世界各国许多杰出的数学家对它进行了研究，并采取了不同的途径，发表了大量的著作，但进展是相当缓慢的，从下面的逐渐趋近最后解决的成果不难看出这一点。

（1）对 a_n 的检验的系列成果是：

①1916 年，比巴霸赫　　证明了 $|a_2| \leqslant 2$；

②1923 年，朗　纳　　证明了 $|a_3| \leqslant 3$；

③1955 年，卡拉贝黛安与希佛　　证明了 $|a_4| \leqslant 4$；

④1968 年，佩德森与奥赞韦　　证明了 $|a_6| \leqslant 6$；

⑤1976 年，佩德森与希佛　　证明了 $|a_5| \leqslant 5$；

（2）先证 $|a_n| \leqslant k_n$（$n = 1, 2, \cdots$），k 为常数，然后逐步改进 k 并使它趋于 1，其成果进展是：

①1916 年，比巴霸赫　　证明了$|a_n|<\frac{e^2}{4}n$；

②1925 年，利特尔伍德　　证明了$|a_n|\leqslant en$；

③1945 年，罗鲁金　　证明了$|a_n|<\frac{3}{4}en$；

④1965 年，法基列维克　　证明了$|a_n|\leqslant 1.3586n+1.51$；

⑤1965 年，米　林　　证明了$|a_n|<1.243n$；

⑥1972 年，菲尔杰拉尔德　　证明了$|a_n|<1.081n$；

⑦1978 年，霍罗威茨　　证明了$|a_n|<1.0657n$；

(3) 对 s 族中的子族进行研究的成果是：

①设 $s_1=\{f(z);f(z)\in s,f(z)=\sum_{n=1}^{\infty}a_nz^n$ 为实数$\}$ 则当 $f(z)\in s_1$ 时，有$|a_n|\leqslant n$ ($n=1$，2，…)；

②设原点 $z=0\in D$，如 D 中任一点子与原点的连线属于D，则称 D 为关于原点的星形区域。记 $f(|z|<1)$ 为 $|z|<1$ 关于 $f(z)$ 的映象。

设 $s_2=\{f(z);f(z)\in s,f(|z|<1)$ 是关于原点的星形区域$\}$，则当 $f(z)\in s_2$ 时有

$$|a_n|\leqslant n\quad(n=1,2,3,\cdots)。$$

③如果区域 D 内任意两点的连线都属于D，这时称 D 为凸形区域。

设 $s_3=\{f(z),f(z)\in s,f(|z|<1)$ 是凸形区域$\}$，则当 $f(z)\in s_3$ 时，有

$$|a_n|\leqslant n\quad(n=1,2,\cdots)。$$

经过 68 年的相继努力，最后于 1984 年，才由路易斯·德布朗格斯巧妙地利用泛函分析中的算子理论证明了比巴霸赫猜想是正确的。

解决数学猜想不仅需要多种途径，而且也常常经历多次反复才能解决。比如，欧拉方阵猜想自 1782 年提出到 1959 年最后解决，在这 177 年中就至少经历过三次较大的反复。从前面谈到的

可知，第一次反复：1901 年彼得尔森用几何方法证明了 $n=6$ 时，这一猜想是正确的，但 1905 年塔里证明彼得尔森的证明是错误的。第二次反复：1901 年维尔尼克用代数方法证明这一猜想是正确的，但麦克尼许却提出维尔尼克的证明是不对的。第三次反复：1922 年麦克尼许用拓扑方法证明这一猜想是正确的，但 1942 年勒维却证明麦克尼许的证明也是错误的。一直到 1959 年，才由玻色和史里克汉德举出反例，最后彻底否定了欧拉方阵猜想。

又比如，“欧氏第五公设可证”这一猜想，自公元前 3 世纪提出后，世界各国大量数学家投入了这一试证工作，其主要代表人物有：古希腊的波西道尼（公元前 1 世纪），拜占庭的普洛克尔（公元 5 世纪），伊朗的纳西艾丁·屠西（1201—1274），英国的瓦里斯（1616—1703），意大利的萨开里（1667—1773），瑞士的兰佩特（1728—1777），法国的勒让德（1752—1833），俄国的古利耶夫（1764—1813），匈牙利的法·鲍耶，德国的须外卡尔特（1780—1859）等等。他们前赴后继，连续作战，付出巨大的劳动，有的甚至献出毕生的精力，但结果都一一失败了。他们写出大批有关试证欧氏第五公设的书籍和文章，据不完全统计，仅公开发表的这方面专著就有二百五十多部。其中，确实给出了许许多多证明，但仔细考察便可发现，不少人在证明的过程中，自觉或不自觉地引用了与欧氏第五公设相等价的命题，犯了逻辑循环的错误。一直到 19 世纪 20 年代，在总结二千年失败教训的基础上，才获得彻底解决。

研讨数学猜想，有时还要创造出各种新方法才能取得进展。前面讲到的“相邻平方数之间至少存在二个素数”，即杰波夫猜想的研究情况就是如此。为了探讨素数分布的状况，长期以来数学家们就在研究两个相邻平方数之间是否存在素数，存在多少素数等问题。但是，由于素数有无限多个，所以对素数充分大时的情况就难以讨论，研究进展是十分艰难而缓慢的。我们知道杰波

夫猜想亦称“相继素数差猜想”，这是因为“相邻平方数之间至少存在二个素数”可变换为“相继素数差”问题的讨论。事实上，我们用 x，Y 表示实数，考虑区间 $[x, x+Y]$ 是否存在素数，如果令 $x=n^2$，$Y=2n+1$，并说在此区间内至少存在两个素数，那么此问题就变成为杰波夫猜想了。当 x 取为第 n 个素数 P_n 时，若在区间 $[x, x+Y]$ 中存在素数，则第 $n+1$ 个素数 P_{n+1} 必属于此区间，亦即有

$$x<P_{n+1}<x+Y=P_n+Y,$$

所以

$$P_{n+1}-P_n<Y。\qquad (*)$$

现在的问题是：Y 取何值时在区间 $[x, x+Y]$ 中一定存在素数?

我们知道，在自然数列中，可找到连续 K 个合数，对这样的 K，必有 $P_{n+1}-P_n \geqslant K$，这里的 K 可为任意大。可见，如果取 Y 为一个常数，则它无论为多么大，(*) 式都不可能成立。那么，Y 值到底应如何选取呢? Y 可取为 x 的增函数。如取 $Y=x$，即讨论在 $[x, 2x]$ 区间中是否存在素数的问题。不少数学家对此问题进行了研究，并取得一些成果。1845 年，伯特兰 (J.Bertrand) 证明了当 $6<n<6000000$ 时，在 $n/2$ 与 $n-2$ 之间至少有一个素数，此结论后人称之为“伯特兰猜想”。1854 年，切比雪夫用初等方法证明了伯特兰猜想，且证明了，对于每一个大于 1/5 的数 ε 而言一定存在数 ξ，使得每一个大于或等于 ξ 的 x，则“在 x 与 $(1+\xi)x$ 之间至少存在一素数”这一结论是正确的。1888 年，谢尔凡斯脱 (T.J.Selvester) 将 ε 改进为 0.16688。1891 年，斯蒂尔克斯 (T.J.Stieltjes) 和凯恩 (E.Cahen) 证明 ε 可取任意小的正数。从讨论中可知，(*) 式中 Y 的选取关键在于 Y 与 x 的相对大小，我们可用 $Y=x^{\theta}$ 来表示。当 $\theta=1$ 时，就是伯兰特猜想。这样，讨论“相继两平方数 n^2 与 $(n+1)^2$ 之间存在素数”与证明“$\theta=\frac{1}{2}$ 时区间 $[x, x+$

x^{θ}] 中存在素数”是很相近的。1905 年，马伦特（Maillet）证明了至少有一个素数存在于两个小于 $9\cdot10^6$ 的相继平方数之间，但这只是一个个别的事实。究竟所有相继平方数之间是否都存在素数？尚待解决，问题的关键在于考虑 x 充分大时，θ 取何值可使区间 $[x, x+x^{\theta}]$ 中有素数。当个问题，当 $\theta<1$ 时，用初等方法研究十分困难，故迫使人们用新的方法来取得进展。

为此，人们开始用解析方法来加以研究。用 $\pi(x)$ 表示不超过 x 的素数个数，则在区间 $[x,\ x+Y]$ 中的素数个数便可表示为 $\pi(x+Y)-\pi(x)$。我们可以用复变函数论的方法将此式的讨论转化为对黎曼 ξ 函数 $\xi(s)$ 的一些性质的研究。前面已经讲过，1921 年，克拉莫证明了在黎曼猜想正确的前提下，当 $x=P_n$，$Y=P_n{}^{1/2}\log P_n$ 时，则区间 $[x, x+Y]$ 中必定有素数存在。1930 年，豪海塞尔（G. Hoheisel）证明了，当 x 充分大时，一定存在一个小于 1 的 θ，使得在区间 $[x, x+x^{\theta}]$ 中必有素数，并指明 $\theta=32999/33000$ 就可以了。1933 年，海伯伦（H. Heilbronn）证明了 θ 可以取 249/250。1936 年，契达柯夫（H. P. Zygakof）证明了 θ 可以取比 3/4 的任何数。1937 年，英格汉姆（A. E. Iugham）证明了，若存在常数 M 与 C，使得对所有的实数 t，下面不等式成立

$$\left|\xi\left(\frac{1}{2}+t\right)\right|\leqslant Mt^{c},$$

那么，θ 就可以取任意一个比 $\frac{1+4c}{2+4c}$ 小的数。1949 年，闵嗣鹤证明了 c 可以取比 15/92 大的任何一个实数，从而 θ 可以取大于 38/61 的任意实数。1973 年，赫克斯雷（M. N. Hnxley）通过迂回的途径证明了 θ 可以取大于 7/12 的数。至此，用解析方法再没有取得较大的进展。这样，人们便开始考虑改进解析方法，再作进一步的探讨。

后来、人们采取了解析方法与筛法相结合的方法，对相继素

数差猜想进行研究，结果又取得一些进展。1979 年，因凡涅斯（H.Iwaniec）与裘悌拉（M.Jutila），把解析方法与筛法结合起来，证明了当 θ 大于 13/23 时，区间 $[x, x+x^{\theta}]$ 中定有素数。同年，希恩-布朗（D.R.Heath-Brown）和因凡涅斯证明了，当 θ 大于 11/20 时，区间 $[x, x+x^{\theta}]$ 中定有素数。1981 年，楼世拓、姚琦证明了 θ 可取大于 35/64；同年，因凡涅斯证明了 θ 可取大于 17/31；1983 年因凡涅斯与品茨（J.Pintz）证明了 θ 可取大于 23/42；1984 年，楼世拓、姚琦证明了 θ 可取大于 6/11 等等。

从上述这一历史过程，我们不难发现，像数学猜想这样的数学难题，它的研究进展直至最后解决，与改进研究方法有着极为密切的关系。方法创新了，研究就取得新成果。而方法的创新，又是一件很不容易的事。这正是数学猜想艰难性的一种重要表现。

此外，数学猜想的艰难性，也还表现在一些重大的数学猜想，虽出自杰出数学家"数学天才"之手，也有时是错误的。比如，19 世纪欧洲数学权威高斯曾提出一个猜想：素数分布函数 $\pi(x)$ 与对数积分函数 $Li(x)$ 之差为定号，即 $Li(x)-\pi(x)>0$。然而，后来列特虎得（Littlewod）却于 1914 年，证得了：当 x 充分大后，$Li(x)$ 与 $\pi(x)$ 之差的符号无限次改变。这样，便以这个事实否定了高斯的上述猜想。此外，前面提到的费尔马猜想、欧拉方阵猜想等，虽都是像费尔马、欧拉这样的数学权威提出来的，但结果却被后来人们所否定。由此可见，再优秀的数学家所提出的猜想也可能是错误的。

（三）有时得不到多数人的承认

数学猜想的艰难性，不仅表现在提出和研讨之中，而且也表现在取得研究成果后，有时得不到多数人的承认，甚至遭到一些

人的排斥、反对和攻击。这里，我们列举数学发展史上的二个事例，来具体说明这一点。

其一，作为“欧氏第五公设可证”这一数学猜想研究成果的非欧几何理论，曾遭到传统观念的强烈抵制。

由于非欧几何无论是理论本身还是建立理论的途径，都与传统观念格格不入，而罗巴切夫斯基运用反证法论证平行理论的技巧又是相当高超的，思想是极其深刻的，致使当时许多数学家不理解；加之他所提出的这些新理论，当时没有在实践中检验和应用，因此，在相当长的一段时间里，未得到多数人的承认，甚至遭到了传统观念的非难和阻挠。1826 年，罗巴切夫斯基在俄国喀山大学宣读了他关于非欧几何的研究报告，当时该校的学术委员会根本不重视，不仅不予以认真审查和发表，而且还把论文原稿给弄丢了。1829 年，他在《喀山通报》上发表了关于非欧几何的论文。这时，俄国科学院委托当时著名数学家奥斯特洛格拉德斯基来审查这篇论文，可是，这位数学权威根本不理解罗巴切夫斯基的新几何理论，认为罗氏的著作“不值得科学院注意”，因为它的“内容平庸”，“压根儿就是错误的”。《祖国之子》杂志还特意发表了一篇诽谤罗巴切夫斯基的文章，嘲笑他“把荒谬当作重要的发现”，“建立起海涩的、不可解的和神秘莫测的学说”。至于高斯，虽然较早地致力于这个问题的研究，并用反证法推得一系列定理，但由于他屈服于传统观念的压力，怕引起“愚人的叫喊”，终生未敢发表他关于非欧几何的研究成果。而亚·鲍耶终生从事于非欧几何的探讨，也遭到了以他父亲法·鲍耶为代表的传统势力的百般阻拦。法·鲍耶是匈牙利有名的数学家，他由于试证第五公设的失败，变得非常保守。当他听说正在大学学习的儿子亚·鲍耶要试证第五公设时，立即写信说，他熟知一切方法直到尽头，但没有一个是成功的。这一工作埋没了他人生的一切光亮和欢乐。并且深有感触地劝阻说：“老天啊，希望你放弃这个问题，因为它也会剥夺你生活的一切时间、健康、休息，一切

幸福……”。[1] 然而，亚·鲍耶并没有因此停止他的研究工作。1825年，他写出了关于非欧几何的论文，多次要求发表均被他父亲压下去了。后来，他把论文寄给了他的母校维也纳军事工程学院某数学教授，但遗憾的是竟被遗失了。1831年，在亚·鲍耶再三请求下，法·鲍耶同意在他的著作后面以“附录”的形式发表。法·鲍耶把这部著作寄给高斯，而高斯看了亚·鲍耶的论文后，一方面感到惊奇，另一方面回信表示“不赞成亚·鲍耶的工作”。因此，法·鲍耶更加反对亚·鲍耶的这项工作。最后，亚·鲍耶被其父驱逐出家，过着十分贫寒的生活。

直到1854年，黎曼以“关于几何基础上所用的假设”为题发表了演讲，把非欧几何向前推进了一步，但仍未得到公认。1868年，意大利数学家贝尔特拉米在《非欧几何解释的尝试》中，证明了非欧几何可以在欧氏空间的拟球曲面上得到解释，从而它的实际意义得到了间接的说明，非欧几何的思想开始被人们所接受。特别是1905年以后，非欧几何在爱因斯坦相对论、原子物理学及天文观测中得到检验和应用，人们才确认它是现实世界中空间形式的一种正确反映，是科学的几何理论。这一事实充分表明，无论是一次数学方法上的重要突破，还是一个新数学理论的发现和最后确立，都不是一帆风顺的。

其二，对比巴霸赫猜想获得证明的惊人成果，在开始时许多人持怀疑态度，作出证明的美国数学家德布朗格斯却不得不从苏联寻找知音。德布朗格斯在四十多年前曾发表过错误的证明，从那时起，数学家们对他一直很冷淡，不信任。他为证明比巴霸赫猜想，穷想苦战了7年之久。他本人在回顾证明过程时说，这是一个艰难的历程，得不到资助，并受到严重的攻击。1983年5月，他获得了成功，但他无法在美国数学界中找到一名数学家愿意看他那350多页的文稿。他曾先后将自己的文稿至少寄给了十

① 转引自B·И·科士青：《几何学基础》，商务印书馆，1957版，第37页。

二位数学家，但绝大多数人没有看它。据说有一位数学家开始认真看他的文稿，但当发现有错误时，就不再往下看了，其实这个“小错误”不影响证明结果的正确性。在美国得不到支持的情况下，恰巧他有个访问苏联的机会。在访问期间，他在列宁格勒大学向苏联的数学家们报告了他的研究成果。开始时，苏联的数学家也不抱有更大乐观态度，但还能认真地听取他的报告。报告共五次，每次都由下午 5 时到晚上 9 时，有时还更晚。经仔细审查，苏联数学家肯定了他的证明是正确的。后不久，德布朗格斯向苏联斯捷克洛夫数学研究所所长、数学权威杂志主编 L. D. Faddeev 提交了 12 页的预印本。随后，苏联向全世界数学家寄发了此预印本。当美国得知从苏联传出来的德布朗格斯解决比巴霸赫这个世界难题消息之后，一些报刊如《纽约时报》、《科学美国人》、《科学杂志》等作了醒目的报道。许多科学家高度评价德布朗格斯的成就，说他作出了“一项伟大的成果”，“创造了惊人的奇迹”!

九、数学猜想的类型、特征与意义

从前面的讨论中，我们可以看到，数学猜想存在多种多样的表现形式，并有着许多鲜明的特征。为了更深入地理解数学猜想的思想实质，根据我们所掌握的历史资料，现就它的类型与特征，作些初步的分析。

（一）数学猜想的类型

1. 存在型猜想

所谓存在型猜想，即指内容是讨论存在性问题的那些数学猜想。这一类型的数学猜想，按其内容又可分为两种：

（1）只讨论存在与否。比如，“克拉莫猜想”：当 $x=P_n$，$y=P_n{}^{1/2}\log P_n$ 时，在区间 $[x,x+y]$ 中必定有素数存在；“连续统假设”：在“可数”的势与“连续统”的势之间没有其它的势；“欧拉方阵猜想”：半偶数的方阵是不存在的等。

（2）既讨论存在与否，又指明其内容或量的关系。比如，“波文猜想”：方程 $1^n+2^n+\cdots+m^n=(m+1)^n$，只有正整数解 $n=1$，$m=2$ 与“商高数猜想”：对于正整数 a，b，c，x，y，z，如果有 $a^2+b^2=c^2$，则 $x=y=z=2$。其中，不但肯定

方程存在正整数解，而且指明了解就是 $n=1$，$m=2$ 与 $x=y=z=2$。又比如，“巴切特猜想”：对 $n=1$，2，3，…，方程 $n^2=x^2+y^2+z^n+\omega^2$ 至少有一组 x，y，z，ω 的非负整数解；“伯特兰猜想”：在 $n/2$ 与 $n-2$（$6<n$）之间至少有一个素数；“孪生素数猜想”：孪生素数有无穷多等。其中，不仅肯定存在，而且还指明了存在的数量。

2. 规律型猜想

所谓规律型猜想，即指内容是揭示规律性的那些数学猜想。这一类型的数学猜想，按其内容又可分为三种：

（1）揭示某种性质。比如，“费尔马猜想”：当 n 为自然数时，形如 $F(n)=2^{2^n}+1$ 的数均为素数；“泰特猜想”：任何 3-连通的三次平面图都是哈密尔顿的；“唯一分解猜想”：仅当 $D=1$，2，3，7，11，19，43，67，163 时，$a+b\sqrt{-D}$（a，b，D 为整数，$D>0$）可唯一分解为一些素数乘积等。

（2）揭示状态特征。比如，“彭加勒猜想”：在 n 维空间中的一点集，若是 $n-1$ 连通的紧致流形，则必定是 n 维球；“场站设置猜想”：在平面上 n 点连线总长度最短时，其连线间的结点角皆不小于 120°；“黎曼猜想”：函数 $\xi(S)=\frac{1}{1^S}+\frac{1}{2^S}+\frac{1}{3^S}+\cdots$（其中 $S=\sigma+ti$ 为复数）的零点全部落在复平面的一条直线 $\sigma=\frac{1}{2}$ 上等，

（3）揭示量的关系。比如：“比巴霸赫猜想”：若函数 $f(x)=x+\sum_{n=2}^{\infty}a_nx^n$（$x$ 为复变数）在其定义域单位圆内（$|x|<1$）单值、连续，且当 $x=0$ 时，有 $f(0)=0, f'(0)=1$，则 $|a_n|\leqslant n$ 和“牛曼猜想”：任意 n 个连续整数 $m+1$，$m+2$，…，$m+n$ 总可以重新排列成 $m+i_1$，…，$m+i_n$，使 $(m+i_j, j)=1$（j

$=1, 2, \cdots, n$)，就分别揭示了系数与指数、整数 $m+i_j$ 与自然数 j 的关系。

3. 方法型猜想

所谓方法型猜想，即指内容是阐述解决问题的方法与途径的那些数学猜想。这种猜想大都产生于实际工作的需要之中。比如，1950年，作为全国工业基地的东北地区的交通运输十分繁忙。为了合理运输，节约运力，当时东北计委会一个专业运输小组，通过学习与钻研，在进行实际调运方案的制订中，发现了如下规律：对于只有一个环状的铁路线，其上只有两个发点及若干个收点的情况，每一发点向两边供应的最大界限（即分界线）的定法应该满足：顺时针方向的运输线长之和等于逆时针方向之运输线长之和。

$$\overset{\frown}{AC}+\overset{\frown}{BD}=\overset{\frown}{AD}+\overset{\frown}{BC}=\frac{1}{2}\text{全圈长}$$

其中，C 和 D 是分界点。

实际问题往往多于两个发点。如果有多个发点和多个收点的情况又会怎样呢？粮食部的同志在上述特殊情况的基础上，概括出一般的方法，并称之为"图上作业法"。所谓图上作业法，就是在交通图上进行的吨公里数（这里假定重量单位为吨，距离单位为公里）为最小。使之吨公里数最小的运行方案就叫做"最好方案"。为了叙述图上作业法，我们先介绍几个基本概念。

对流：在一段路线上同时有两个相反方向的调运；

流向图：在编造调运方案之前，先给出它的交通图，并用两种记号标出收、发点，把需要调进与调出的数字以及相邻两点间的距离数都在旁边标出来。将发点调往收点的物资数，在路线的右侧按前进方向绘上一个流向。当收发点间需要调出和调进平衡时，就得到一个流向图。

在流向图中，有些流向是在圈的外面，也有些流向是在圈的

里面。在圈外面的流向称为“外圈流向”，在圈里面的流向称为“内圈流向”。各个外圈流向长之和叫做“外圈之长”，各个内圈流向长之和叫做“内圈之长”。

对于一个圈的交通图，它的流向图所确定的调运方案是最好的充分必要条件是：它的流向图上没有对流，而且内圈之长和外圈之长都不大于整个圈长的一半。

在上述结论未获理论证明之前，就运用它来确定调运方案，并认定是最好方案，显然这一做法是具有经验性质的，或者说仅仅是一个“猜想”。1958年，中国科学院数学研究所运筹学研究室的同志们，终于从理论上严格证明了按上述方法所规定的调运方案一定是最好方案。这样一来，图上作业法就成为完全可靠的科学方法了。

此外，在最优化的研究中，出现了各种各样的算法，其中有些算法在相当长的时间内给不出理论上的证明。在没有给出理论证明之前的那些算法，实质上都是“猜想”。

（二）数学猜想的特征

关于数学猜想的特征，在本书的第八部分中，已经详细地分析了它的艰难性。这里，我们再来讨论它的其它特征。

1 真伪的待定性

数学猜想既然是根据某些已知事实材料与数学知识，对未知量及其关系所作出的一种预测性的推断，那么它必然具有两个显著的特点：一是具有一定科学性，二是具有某种假定性。正是由于这两个特点，便决定了数学猜想是处于孕育阶段的、尚待证实和公认的科学思想，也就是说它必然表现出真伪的待定性，究其结果可能被肯定，也可能被否定，还可能是不可判定的。比如，1926年，德国数学家范德瓦尔登（Van der Waerden）提出猜想：

设 A 是 $n\times n$ 矩阵，矩阵元为 a_{ij}（$i=1$，2，…，n；$j=1$，2，…n），则 A 的正项行列式（Permanent）Per（A）定义为

$$\mathrm{Per}(A)=\sum_{\sigma\in S_n}a_{1,\sigma}(1)a_{2,\sigma}(2)\cdots a_{n,\sigma}(n)$$

其中，S_n 表示 n 个符号的对称群。这一猜想，引起了不少数学家的兴趣，并对此进行了探讨与研究。但 50 多年来没有取得突破性的进展，直到 1979 年与 1980 年，苏联数学家费里克曼和埃戈伊夫证明是正确的，从而肯定了这个猜想。这是被肯定的。还有被否定的，比如，前面讲到的“费尔马猜想”、“欧拉方阵猜想”与“泰特猜想”等，最后都被否定了。

数学猜想中可能被肯定和否定的以外，还有不可判定的。比如，前面提到的“连续统假设”就属于这一类。事实上，1882 年，康托尔提出：在“可数”的势与“连续统”的势之间没有其它的势。1938 年，哥德尔证明了由 ZFC 公理系统推不出连续统假设的否定式。1963 年，柯恩又证明了 ZFC 公理系统推不出连续统假设。这两个结论即证明了连续统假设在 ZFC 公理系统中是不可判定的命题。就是说，连续统假设在 ZFC 公理系统中，既无法判定是真，亦无法判定是假，只有开辟新的领域才能有所解决。

2. 思想的创新性

数学猜想既然是对未知量及其关系的一种预测性的推断，又常常是科学理论的萌芽和胚胎，那么它的思想必然是具有创新性的。“创新”是数学猜想的灵魂，没有创新就没有数学猜想。这种创新首先表现在提出新的见解上面。比如，“欧氏第五公设可证”这一数学猜想，就提出了与《几何原本》不同的新观点。欧氏第五公设在《几何原本》中是作为演绎推理的前提而出现的，认为它不需要也不可能作出数学证明。但是，上述数学猜想恰恰相反，认为它是可以给出数学证明的，不是前提，而是结论；不

是公理，而是定理。这显然是一种新的见解，也正因为如此，便引起了许多数学家的兴趣，进行了大量的试证工作，并导致了非欧几何这一新几何分支的诞生。

其次，数学猜想的创新性还表现在预见新的事实上面。比如，前面讲过，瑞士著名数学家伯努利对自然数平方的倒数这一无穷级数之和，长期想求得但一直求不出来，深为其艰难而感叹。后来，欧拉对这一难题进行了深入研讨，他通过大胆而巧妙的类比，提出了“$\frac{\pi^2}{6}=1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots$”这一数学猜想。这一猜想即预见了一个新的事实，即自然数平方的倒数这一无穷级数之和等于$\frac{\pi^2}{6}$。后来，从理论上证明了这一预见是正确的。

第三，数学猜想的创新性也表现在揭示新的规律上面。比如，尺规作图是几何学中一个极为重要的问题。在这一问题的探讨中，人们总是力图弄清在几何图形中哪些能够用尺规作出，哪些不能用尺规作出，即揭示和发现其中的规律性。正是适应这种需要，德国数学家高斯提出猜想：所有边数等于费尔马数 $F(n)=2^{2^2}+1$ 中素数的正多边形，均可用尺规作图作出来。这一猜想明确揭示了一些特殊正多边形是可用尺规作出的规律，从而将这一问题的探讨向前推进了一大步。事实上，后来高斯不仅亲自作出了一些符合上述猜想条件的正多边形（如正十七边形），而且还从理论上证明了这一猜想是正确的，肯定了这一规律。

3. 目标的具体性

数学猜想作为一种预测性的推断，必然具有结论即探索目标的具体性。我们知道数学猜想中所给出的结论都是明确、具体的。诸如，“有解”、“无解”、“可证”、“不可证”、“可作”、“不可作”等，但是，一般数学问题并非这样明确、具体。因此，与一般数学问题相比，目标的具体性是数学猜想的一个明显特征。

事实上，无论是哪种类型的数学猜想，都具有这种具体性。我们先来看存在型猜想。这种猜想有时明确提出要求解决的目标是具体的某种对象存在还是不存在。比如，“费尔马大定理”这一数学猜想，就明确指出要解决的具体目标是当 n 为大于零的整数时方程 $x^n+y^n=z^n$ 不存在正整数解。这类猜想也有时不仅具体指出这种存在性，而且还指出存在的具体内容或多少。比如，柯召—孙琦猜想，就一方面指出方程 $x^n+(x+1)^n+\cdots+(x+h)^n=(x+h+1)^n$ 存在解：(1) 当 $n=1$，$h=1$ 时，$x=1$；(2) 当 $n=2$，$h=1$ 时，$x=3$；(3) 当 $n=3$，$h=2$ 时，$x=3$。另一方面，又指出此方程再无其它正整数解。杰波夫猜想则在指出相邻平方数之间存在素数的同时又具体指出至少有 2 个。

其次，我们来观察规律型猜想。它的这种具体性主要表现在指明了要解决的目标是某种具体性质、形态特征或量的关系。像前面讲到的泰特猜想、黎曼猜想和比巴霸猜想这三个数学猜想，就分别反映了上述三种情况。

最后，我们再来分析方法型猜想。这种猜想无不明确指出要解决的目标就是确定有效的具体方法。只要我们仔细地考察一下“图上作业法”、“场站设置猜想”等方法型猜想产生过程和内容，就会清楚地看出它们都是一些为解决实际问题而提出来的可行途径与方法，要解决的目标是明确而具体的。

（三）研讨数学猜想的重要意义

数学猜想是数学研究的一种科学思维形式，是解决数学理论自身矛盾和疑难问题的一个有效途径。研讨它，对数学理论的发展有着极其重要的意义。那么，研讨数学猜想到底有哪些重要意义呢？

1. 丰富数学理论

数学猜想，作为数学的一种潜在形态，作为数学理论的“胚胎”、“萌芽”，它在建立、丰富和发展数学理论的过程中，必然起着“中介”和“桥梁”的作用，其具体表现有以下三个方面：

（1）假若某个数学猜想最后被证明是正确的，那么它就转化为数学理论，从而丰富了数学内容。一般说来，数学猜想被肯定之后，即成为数学定理。这是数学猜想“中介”、“桥梁”作用的最主要的表现。比如，瑞士著名数学家欧拉提出关于超越数的猜想：幂 α^{β}，其中底 α 为代数数，指数 β 是代数无理数，例如数 $\alpha^{\sqrt{2}}$或 $e^{\pi}=i^{-2i}$，总是超越数，至少是个无理数。这一猜想自 18 世纪下半叶提出后，吸引了许多数学家，并力图证明之，但长期没有获得解决。直到 1934 年，苏联数学家盖尔方德和德国数学家施耐德尔，运用抽屉原则，丢番图逼近思想等，分别证明了这一猜想是正确的，并从而转化为数学定理。又比如，数学家莫德尔于 1922 年提出猜想：任一不可约、有理系数的二元多项式，当它的“亏数”大于或等于 2 时，最多只有有限个解。1984 年，德国数学家弗尔廷斯证明了这一猜想是对的，并从而转化为数学定理。

（2）即使某个数学猜想未获最后解决，但在研讨的过程中，却往往创造出一些意想不到的理论成果。比如，自 1889 年提出“黎曼猜想”后，一百多年中，虽然许多数学家付出千辛万苦，但直至今日仍未最后解决。这是问题的一个方面。而另一方面，人们却在探讨这一猜想的过程中，尤其在假定本猜想是正确的基础上，获得了一系列新的重要结论。1901 年，冯·柯赫在假定黎曼猜想成立的前提下，证明了最理想的素数定理的误差各项，即证明了

$$\pi(x)+lix=o(\sqrt{x}\ln x),$$

其中 $\pi(x)$ 表示不超过 x 的素数个数，

$$lix = \lim_{h \to 0}\left(\int_0^{1-h} + \int_{1-h}^{x}\right)\frac{dt}{\ln t}。$$

特别应当指出的是，本来有些结论是在黎曼猜想成立的前提下得到的，但后来为了使证明严格化，却绕过这一猜想得到了最后确定。就是说，黎曼猜想有着发现新理论的功能。事实上，1965 年，数学家朋比利在研讨哥德巴赫猜想过程中，采用“大筛法”证得了：偶数＝（1+3），从而取得了当时这项研究的最好成果。但是，在假定黎曼猜想成立的前提下，这一结果是容易得到的。朋比利的出色之处正是在于他绕开黎曼猜想而证得了这一结果。与当时许多数学家在假定黎曼猜想成立条件下得到不少结论但未获证明相比，朋比利的成就就显得特别突出，因而得到了数学界的高度评价，并荣获了费尔兹奖。

研讨黎曼猜想的重要性还在于：如果它被证明是正确的，那么至今尚存的许多数学难题将获得重大突破，甚至迎刃而解。像前面讨论过的杰波夫猜想，不少数学家为证明它，特将其转化为“当 x 是充分大的实数时，在 x 与 $x+x^{\frac{1}{2}}$ 之间是否一定有素数”这样一个问题，并得到了许许多多成果，其中最好的成果是因凡涅斯与希思一布朗于 1979 年得到的“在 x 与 $x^{11/20}$ 之间一定有素数”这一结论。但是，如果假定黎曼猜想成立，那么便可用极简单的方法证明“在 x 与 $x^{1/2}\log x$ 之间一定有素数”这一结论。

同样，等差级数中最小素数问题的研究也是如此。假定一个等差级数的首项为 a，公差为 d，它的一般项可写为 $a+nd$（$n=1, 2, \cdots$）。我们来讨论，当 n 取多大时，可使 $a+d$，$a+2d$，…，$a+nd$ 中一定有素数？假若有这样一个常数 C，使得等差级数中不超过 d^c 的那些项，即 $\{a+la, l=1, 2, \cdots\}$ 中一定有素数，那么，这里的 C 越小其结果也就越好。1957 年，我国数学家潘承洞最先确定出 $C \leqslant 5448$。1979 年，我国数学家陈景润获得了更进一步的成果，即得到 $C \leqslant 17$。然而，如果黎曼猜想是对的，那么我们便可得到 C 的理想结果：$C \leqslant 2+\varepsilon$，其

中 ε 为任意小的正数。

以上讨论的是，假定数学猜想成立的前提下，获得一些重要结果，其中有的绕开此猜想而获证，也有的等待此猜想最后确证后方能取得突破。下面我们再来分析另外一种情况，即某数学猜想未最后解决，但在寻解过程中，却直接得到一些重要成果，使其成为一个“下金蛋的母鸡”。就拿费尔马大定理来说，为了证明它，三百多年来，许多大数学家欧拉、高斯、狄利克雷、柯西、库默等，都付出过辛勤的劳动，在试证的过程中，得到一些新的成果，像库默创立的理想数论，不仅为建立代数数论这一重要数学分支奠定了基础，而且还成为其它许多数学分支的有效工具。

(3) 虽然某个数学猜想被否定了，但在否定的过程中，却有时发现一些其它方面的数学理论。前面讲到过，“欧氏第五公设可证”这一猜想最后被否定了。但是，它的否命题：“欧氏第五公设不可证”却被证明是正确的，即成为数学理论之内容。特别应当指出的是，在这否定证明获得成功的同时，奇妙地发现并建立了一种崭新的几何理论——非欧几何学，为几何学的发展作出了划时代的贡献。由上可见，研讨数学猜想对丰富数学理论，推动数学发展，具有极其重要的价值。

2. 促进数学方法论的研究

研讨数学猜想的重要意义，不仅表现在它可以丰富数学理论，推动数学科学的发展，而且还表现在它能够促进数学方法论的研究。

(1) 数学猜想作为一种研究方法，它本身就是数学方法论的研究对象。我们在前面已经系统地分析了数学猜想的提出方法，判定途径，以及类型和特征等。这些内容，实际上均属于数学方法规律性问题的探讨。就是说，这里所概括出来的方法、途径、类型和特征等，对总结一般科学方法产生的条件、发展的途径、

形成的类型和呈现的特征，有着重要意义，特别是对创造性思维方法的研究具有特殊的价值。像提出数学猜想的变换条件法与逐级猜想法，判定数学猜想真伪性的逐次趋近与命题转化等，对其它科学方法规律的研究，都有直接参考作用。

（2）研究数学猜想的过程中，又创造许多新方法，从而丰富了数学方法论的研究对象。比如，关于费尔马大定理的证明，费尔马本人解决了 $n=4$ 的情形，实际上他证明了“不定方程 $x^4+y^4=z^2$，$(x, y)=1$，没有 $xy\neq0$ 的整数解”这个更强的结果。在其过程中，他创造了“无穷递降法”。这一方法至今仍发挥着它的作用。费尔马大定理属于不定方程问题。我们知道，不定方程是数论中一个古老而重要的分支，两千多年来，虽然人们取得不少研究成果，但至今尚未发现更有效的一般方法。不久前，英国数学家贝克创造了“有效方法”，对一大类高次不定方程求出了它们正整数解的上界，但对三个变元以上的不定方程，却又无能为力了。又比如，在探讨哥德巴赫猜想的过程中，1930 年，史尼尔曼创造了“密率法”；1973 年，陈景润改进了古老的“筛法”。还比如，人们为解决“连续统假设”这一数学猜想，相继创造了“可构成性方法”与“力迫法”。再比如，在证明“四色猜想”的过程中，创造了具有深远意义的机器证明方法等等。

这里应强调的是，有时数学猜想本身就是数学方法，它的产生自然会丰富数学方法论的研究对象。比如，在运筹学中有这样一个问题：已知平面上有 n 个点，每个点对应一个重量，今在平面上求一点 x，使每个已知点的重量集中在 x 点上的总吨公里数为最小。对这一问题，本世纪 60 年代，波兰数学家鲁卡雪维奇从解非线性方程组出发提出一种计算比较简单的迭代法。他几次实际应用，发现这个方法都是收敛的，但未给出证明。在未给出证明之前，这种方法实质上就是一个“猜想”。现在已经被证明是正确的。

（3）数学猜想作为数学发展的一种重要形式，它又是科学假

说在数学中的具体表现，并深刻反映了数学发展的相对独立性与数学理论相互导出的合理性。恩格斯指出："数学是从人的需要中产生的，……但是，正如同在其他一切思维领域中一样，从现实世界抽象出来的规律，在一定的发展阶段上就和现实世界脱离，并且作为某种独立的东西，作为世界必须适应的外来的规律而与现实世界相对立。"又说："数学上各种数量的明显的相互导出，也并不证明它们的先验的来源，而只是证明它们的合理的相互关系"。[①] 数学猜想是恩格斯上述论述的生动体现。事实上，从前面我们考察与分析的大量事例中，不难看出，数学猜想确是在数学发展到积累了大量资料，需要进行理论整理，探索其理论内部的矛盾规律这一阶段上产生出来的，因而大都表现为命题的逻辑判断形式，并运用思维规律来判定其真伪性。也正是因为数学发展具有相对独立性，数学理论的相互导出具有合理性，所以数学家从数学理论自身的体系中提出一些数学猜想，才有其科学的预见性，可以吸引许许多多数学工作者，而且往往在相当长的时间内还可以成为促进数学发展的中心课题，甚至代表着数学研究的方向。1900 年，德国杰出数学家希尔伯特提出了包括有数学猜想在内的二十三个问题，在一定程度上影响着 20 世纪以来的数学发展。实际上，八十多年来，世界各国的大量数学家被其中的一些著名的数学猜想（如连续统假设、黎曼猜想等）所吸引，进行了大量的研究工作，并且常常把在这些问题的研究上是否有进展作为衡量一个数学家乃至一个国家数学水平的标志之一，作为一种极高的数学荣誉。1976 年美国一些著名数学家评选 1940 年以来数学十大成就，其中有三项就是希尔伯特所提出的第一、第五、第十三问题的解答，第一问题就是著名数学猜想"连续统假设"。伟大科学家牛顿深刻指出："没有大胆的猜测就

① 恩格斯：《反杜林论》，人民出版社，1970 年版，第 35 页。

作不出伟大的发现”。① 这是被大量事实证明了的一条客观真理，遵循这条真理就可以推动数学不断向纵深方面发展。

3. 推动潜科学学的探讨

深入研讨数学猜想，不仅对丰富数学理论，促进数学方法论的研究有积极作用，而且对推动潜科学学这门新学科的理论探讨也有重要意义。以潜科学为研究对象的潜科学学，是本世纪 70 年代末，由我国学者开创的一门新兴综合性学科。它以特有的生命力，顽强地发展着，必将为丰富我国科学学与软科学的研究，创造独具特色的中国人的科学观，不断作出贡献。数学是中华民族擅长的一门学科，在研究和解决数学猜想方面，我国数学家有着光荣的历史。而数学猜想又是科学猜测这种潜科学形态的主要组成部分，因此，系统考察与研究数学猜想及其思想方法，对促进潜科学学理论的发展具有特殊的价值。

(1) 数学猜想属于潜科学形态

数学同其它科学一样，都有一个由“潜”到“显”的发展过程。而数学猜想作为数学发展的一种形式，它是数学发展的“潜”阶段产物，属于孕育中的潜科学形态。

①从数学猜想的待定性看，它具有潜科学特征。关于数学猜想具有待定性，我们前面已经讨论过了。而待定性恰是潜科学区别于显科学的一个基本特征，因此数学猜想属于潜科学范畴。其实，仅就名称而言，既然叫“猜想”，顾名思义，它就具有待定的意思，从而体现“潜在”的性质，“猜”必“潜”，“潜”是“猜”的本质特征。关于素数个数的研究，人们开始能数出小于 100 的素数个数，后来数出小于 1000 的素数个数，小于 10000 的素数个数，……并列出素数个数表。约在 1800 年，高斯与勒让德通过对此表的观察提出猜想，比 x 小的素数个数逼近于 $x/$

---

① 引自贝弗里奇：《科学研究的艺术》，科学出版社，1979 年版，第 153 页。

$\log x$，即 $\pi(x) \sim x/\log x$。从这一猜想提出的过程，我们不难看出，一方面，它是以素数个数表中的具体数据为依据而提出的，故具有一定科学性；另一方面，猜想中又包含表中未涉及的数（小于此数的素数个数），加之未有理论证明，故又具有其假定性。就是说，这一猜想是一种孕育中的数学思想，具有待定性。约在 1850 年，切比雪夫对这一猜想的研究取得部分结果，作出重大贡献。1896 年，阿达马和普辛用复变函数论的方法，各自独立地证明了这一猜想是正确的，并从而获得了著名的“素数定理”。

②从数学猜想的创新性看，它具有潜科学价值。我们知道，数学猜想的一个突出特征是它的创新性，即具有提出新见解、预见新事实、揭示新规律的作用。这些作用正是潜科学的价值，因为科学的生命在于创新，在于革命，科学由“潜”到“显”的转化过程，即是创新、革命和发展的过程，也只有体现这一过程的潜科学，才真正称得上潜科学。数论这一重要数学分支，可以说就是在提出和试证一系列猜想的过程中，不断向前发展的。像古希腊时提出“四平方猜想”：每个正整数都是四个或四个以下平方数之和，后来人们证明它是对的，故转化为“四平方定理”；19 世纪初，提出“高斯—勒让德猜想”，19 世纪末，阿达马和普彦证明这一猜想是正确的，从而获得素数定理等，这样的事例在数论发展历史上是屡见不鲜的。尤其是关于素数的分布和不定方程解这两个数论的古老问题，一直充满着数学猜想，有些解决了，也有些至今尚未解决，并且还在不断提出新的猜想。

数学猜想的潜科学价值，还表现在研讨它的过程中有时可以创造出具有普遍意义的数学方法。比如，美国数学家阿佩尔和黑肯在解决四色猜想过程中，开辟了机器证明方法应用的广阔领域。这种方法不仅解决了四色猜想，而且更为解决其它一些数学问题，提供了有效的研究方法，具有深远的意义。有人认为，这是数学思想方法上的一次重大变革。因为过去人们从未想到机器

可以解决纯理论性质高难度的定理证明的问题，或者说，从未想到机器可以代替这种创造性思维活动，而今天却成为现实，所以它是数学发展史上的一个重大事件。十多年来，我国著名数学家吴文俊，先后在初等几何定理的机器证明和微分几何定理的机器证明上，取得突破性的进展，得到了国际数学界的好评。

③从数学猜想的艰难性看，它是更深层次的潜科学思想。一般说来，数学猜想是数学上的重大难题，不是短时间可以解决的。这在一定意义上说明它是更深层次的潜科学思想。而具体反映这一点的是，有的数学猜想是在假定另一个数学猜想成立的前提下提出来的，如果说假定成立的数学猜想是第一个层次的潜科学思想，那么以这假定成立的数学猜想为条件而提出来的数学猜想，就是第二个或更深层次的潜科学思想了。像前面讨论过的 n 生素数猜想、连续统假设的 82 个推论和克拉莫猜想等，都属于此类。

(2) 丰富潜科学学研究的内容

①从中探讨潜科学形态产生与演进的规律

潜科学形态包括：科学问题、科学幻想、科学猜测、科学悖论、科学经验与科学蒙难等。数学猜想是科学猜测在数学中的具体表现。我们深入研究数学猜想，考察与分析它提出与发展的过程，必然有益于把握科学猜测这种潜科学形态产生与演进的规律。前面我们已经系统地讨论了数学猜想提出的方法和判定其真伪性的途径，其中有些方法、途径是具有普遍意义的，如类比法、模拟法、观察法、举例否定、命题转化、机器证明等，无论对物理学、化学中的猜测，还是对天文学、地学、生物学中的猜测，都有重要参考价值。这是从方法上分析产生与演进，如果从原因上探讨其发生与发展，也会发现一些带有共性的东西。比如，一般说来，数学猜想发端于数学问题，产生于数学理论体系自身的矛盾之中，这显然对探讨其它学科中猜测产生与发展规律具有指导意义。

②从中探讨潜科学的基本特征

我们知道，潜科学具有待定性、隐变性、创造性、高难性、趋显性等基本特征。显然，数学猜想的待定性、创新性、艰难性等特征是潜科学特征的具体表现。其实，潜科学的隐变性，在数学猜想中也有一定程度的体现。孕育中的数学思想，常常表现为人们头脑中的潜意识流，因而必然有着隐变性的特征。这种隐变性，主要表现在数学猜想的提出有一个忽隐忽现，若暗若明，由模糊到清晰，由不完善到完善的过程。比如，前面讨论过的“一切凸多面体的面数（F）、顶点数（V）与棱数（E）的关系为 $F+V=E+2$”这一猜想，从提出来看有一个酝酿、萌发、修改与完善的过程，或者说，它是经过一系列的隐变过程后才提出来的。至于潜科学的趋显性，在数学猜想中亦有反映。从数学发展史上看，许多数学定理是由数学猜想转化而来的。为什么会这样呢？就是因为数学猜想具有趋显性。也正是因为如此，人们才千方百计去发现与解决数学猜想，以达到发展数学理论的目的。这是从形式上看。从实质上看，因为数学猜想是根据一些已有的数学知识与事实而提出来的，即有一定科学依据的，因而这些科学依据就自然成为它具有趋显性即转化为数学理论的根据了。不仅如此，我们在探讨数学猜想基本特征产生原因时，还可能发现一些有普遍意义的东西。比如，深入分析数学猜想艰难性产生原因，就会发现，传统观念束缚、学术权威压制、思想方法片面等是造成研究成果得不到多数人承认的主要原因。而这些原因又常常是其它科学成果遭到埋没的主要因素。

③从中探讨科学由“潜”到“显”的转化机制。科学由“潜”到“显”的转化是科学发展的一条客观规律，其转化机制表现在许多方面，诸如，提高科学素养，树立正确的科学成败观，开展自由论争，加强管理，倡导伯乐精神等等。从前面讨论中，我们可以看出，数学猜想从提出到解决，从“潜”到“显”，也都与这些机制有关。因此，研究数学猜想将有助于把握科学由

"潜"到"显"的转化机制。这里我们着重就开展自由论争方面的问题加以讨论。围绕数学猜想展开论争，并由此获得其彻底解决，这在历史上是常见的一种现象。比如，狄利克雷原理这一数学猜想自提出后，经历了三十多年的激烈论争和反复，最后才确立起来。所谓狄利克雷原理，是指德国数学家狄利克雷在研究微分方程位势原理时提出的一个猜想，其具体内容大体是：极小化狄利克雷积分

$$\iint \left\{ \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right\} dxdy$$

的函数 u，满足位势方程

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0。$$

后来有人在研究三维位势方程（亦称拉普拉斯方程或调合方程）：

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

时，又提出，由位势方程所描述的相应物理状态总有一个确定的物理解，因而其本身也必然存在一个数学解。但在数学上的这种存在性，迟迟得不到证明。1851 年，黎曼在他的博士论文《单复变函数一般理论的基础》中，给出了位势方程边界问题解的存在性证明。由于黎曼在文中运用了他的老师狄利克雷所提出的上述猜想，故他称之为"狄利克雷原理"。可是，事隔不久，便引起了热烈论争，特别是黎曼的这一证明遭到了德国著名数学家魏尔斯特拉斯（K.W.Weierstrass，1815—1897）的尖锐批评，说黎曼不加证明就先验地假定存在一个使积分达到极小值的函数，这在数学上是不允许的。黎曼没有因此动摇自己对狄利克雷原理的信任，并运用此原理又作出一系列重要发现。1866 年，黎曼去世了，但论争仍未停止。1870 年，魏尔斯特拉斯给出了一个与狄利克雷原理相反的例子。在这个例子中，对给定的边界条件，使狄利克雷积分达到极小值的函数是不存在的，并以此来否定狄利克雷原理。由于狄利克雷原理遭到了数学权威魏尔斯特拉

斯的否定，故数学家们只好另寻他径来证明位势方程边界问题解的存在性，1870年，牛曼用“算术平均值法”重新给出一个证明；1890年，施瓦尔茨利用“交替法”又给出一个证明，同年，彭加勒用“扫散法”也给出一个证明等等。这些证明显然在逻辑上是通得过的，但没有一个能像以狄利克雷原理为工具证明那样简单、明快。许多数学家见此情景，越发对“狄利克雷原理”的被否定感到惋惜，并由之产生复活这一原理的念头，作出一些努力，然而未能成功。数学家牛曼悲观地表示：如此美而又有如此广阔应用前景的狄利克雷原理，已经从我们的视线中“永远消失”掉了！

就是在狄利克雷原理被否定三十年之后，即1899年，德国另一位著名数学家希尔伯特，与魏尔斯特拉斯又展开了论争。他冲破那种把严格性与简单性对立起来的传统观念，批评魏尔斯特拉斯以严格性全盘否定狄利克雷原理的作法，从狄利克雷原理的简单性、优美性以及应用的有效性出发，积极寻求它的真实性和合理性，最后终于找到了证明狄利克雷原理的途径和方法。他的德国数学联合会上报告了他的这一研究成果，并明确指出：只要对问题中的区域、边界值和允许函数的性质作适当的限制，就可以恢复狄利克雷原理的真实性。他还针对数学家们认为狄利克雷原理早已沉没了的观点，意味深长地将他的这一研究工作称为“狄利克雷原理的复活”。① 后来希尔伯特又给出一个更为一般的证明，从而进一步肯定了狄利克雷原理存在的合理性。从狄利克雷原理死而复生的历史事实中，我们不难看出，自由论争对数学猜想的研究和解决，有着积极的推动作用；同时，要想使真正、合理的数学猜想不被埋没，并充分发挥它应有的作用，还必须有远见卓识的数学家为之奋斗，并作出真正科学的理论判定。这些历史经验，无疑对深入认识科学由“潜”到“显”的转化规律是

---

① 参见赵树智：《狄里克雷原理的沉浮》，《潜科学杂志》、1985年第二期。

大有益处的。

要想使潜科学研究成果不被埋没，还有许多条件。像成果的审查者要有无限的耐心和责任感，不要因存在一些微不足道的缺欠而否定其根本的正确性，就是一个重要因素。比如，前面提到的比巴霸赫猜想，当路易斯·德布朗格斯经过 7 年的艰苦努力宣布这一猜想彻底解决时，许多数学家持怀疑态度，尤其在美国几乎找不到一个支持者。究其原因除了这个猜想难度大，过去不少人说解决了实际是错误的以外，还有一个原因就是审查者不耐心。当时，德布朗格斯将他的 350 多页的论文送给一些数学家审查时，美国有一位曾审查过多篇关于这个猜想证明论文（结果都错）的著名数学家，开始还能认真地看，但当发现其中有错误时，他就再不往下看了。其实，这里的错误并不防碍最终结果的正确性。后来，德布朗格斯只好在苏联找到了耐心者，最后得到肯定。类似的情况，在其他学科领域里也是常发生的，故具有普遍意义。

数学猜想一览表

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
欧氏第五公设猜想	欧氏第五公设可证			公元前三世纪	直观推断	德国 俄国 匈牙利	高斯 罗巴切夫斯基 亚·鲍耶	19世纪20年代	反证法	否定 创立非欧几何理论	使几何学发生了一次革命	
费尔马大定理	当 n 为大于2的整数时，方程 $x^n+y^n=z^n$ 没有正整数解	法国	费尔马	1637	不完全归纳法	英国	安德鲁·维尔斯	1994	转化	肯定	试证中创立了“理想数论”等新分支	
默森尼猜想	当 p 为素数时，形如 $M(p)=2^p-1$ 的数中有无限多个素数		默森尼	1644	不完全归纳法			1979		未定 发现27个这样的素数 1979年，电子计算机算出 $2^{44497}-1$ 是素数，有一万三千多位		
费尔马猜想	当 n 为自然数时，形如 $F(n)=2^{2^n}+1$ 的数均为素数	法国	费尔马	1664	不完全归纳法		欧拉	1732	举反例	否定		

续表 1

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
哥德巴赫猜想	每个大于 4 的整数均可表示为两个素数之和	德国	哥德巴赫	1742	不完全归纳法	中国	陈景润	1973	逐次趋近	未定 证明偶数 = (1 + 2)		证明偶数 = (1 + 1)为最后解决
华林猜想	任一正整数必为 4 个平方数，9 个立方数，19 个四次方数之和。(对任意给定的正整数 n，是否存在一个 $r = r(n)$，使得对任意正整数 N，不定方程 $N = x_1^n + x_2^n + \cdots + x_r^n$ 恒有解，$x_i \geqslant 0$ 为整数。)		华林	1770		德国	希尔伯特， 哈代	1909 1919		证明一般形式的，即括号内的。 给出数 $r = r(n)$ 的渐近公式	1934 年，苏联维诺格拉朵夫创造了使 $G(n)$ 估值急速下降法	此猜想发表于《代数沉思录》
欧拉方阵猜想	半偶数的方阵是不存在的	瑞士	欧拉	1782	不完全归纳法	印度	玻色 史里克汉德	1959	举反例	否定		
欧拉猜想	$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$	瑞士	欧拉		类比	瑞士	欧拉	提出十年后		肯定		

续表 2

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
孪生素数猜想	孪生素数$(P, P+2)$有无穷多				变换条件法					未定 已发现最大孪生素数（$10^{12}+9649$, $10^{12}+9651$）		
三生素数猜想	三生素数有无穷多				变换条件法							
n 生素数猜想	n 生素数有无穷多				逐级猜想							
“唯一分解”猜想	仅当 $D=1,2,3,7,11,19,43,67$ 和 163 时，$a+b\sqrt{-D}$(a、b 与 D 为整数，$D>0$) 可唯一分解为一些素数的乘积	德国	高斯	1797		德国 美国	采格尔 格罗斯	1983		肯定		
高斯一勒让德猜想	比 x 小的素数个数逼近于 $x/\log x$，即 $\pi(x)\sim x/\log x$	德国	高斯 勒让德	1800	观察	法国	阿达马 普辛 赛尔贝尔 爱多士	1896 1949	复杂函数论 初等方法	肯定 肯定	获得重要的素数定理	

续表 3

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
高斯数猜想	高斯引进了一类数(高斯数),并求出了 9 个高斯数,即 1,2,3,7,11,19,43,67,163.他猜想,只有 9 个高斯数	德国	高斯	1800	直观推断	美国	H·斯塔克	1966		肯定		
库姆尔猜想	令 $m \neq 2(\mathrm{mod}4)$, $\xi_m = e^{2\pi i/m}$。h_m 和 h_m^- 分别表示圆域 $K = Q(\xi_m)$ 和它的极大实子域 $K^+ = Q(\xi_m + \xi_m^{-1})$ 的理想类数。$h_m^{-1} = h_m/h_{m^+}$,当 m 为奇素数 p 时,猜想 $h_p^{-1} \sim 2p\left(\frac{p}{4\pi^2}\right)^{\frac{p-1}{4}} = 2p\left(\frac{p}{39.4784}\right)^{\frac{p-1}{4}}$(当 $p \to \infty$ 时)	德国	库姆尔	1850	不完全归纳法	美国	Cavl-itz	1961	逐次逼进	未定		
							Mets-ānky-la	1974		证明了 $h_p^- < (p-1)\left(\frac{p-1}{2}\right)^{\frac{p-1}{4}}$ 证明了 $h_p^- < 2p \cdot \left(\frac{(p-1)(p-2)}{24p}\right)^{\frac{p-1}{4}}$		
						中国	冯克勤	1982		证明了 $h_p^- < 2p\left(\frac{p-1}{31.9971}\right)^{\frac{p-1}{4}}$		
四色猜想	在平面(或球面)上画地图,只要有四种颜色即可保证相邻区域不用同一色	德国	莫比乌斯	1840	经验概括	美国	阿佩尔黑肯	1976	机器证明	肯定	开辟数学研究新途径具有重要认识论意义	

续表 4

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
凯特兰猜想	除 $8=2^3, 9=3^2$ 以外，没有两个连续整数都是正整数乘幂		凯特兰	1842 (1844)	不完全归纳法		卡塞尔斯	约 1961		未定 证明了不存在三个相邻的整数，都是完全幂		
伯特兰猜想	在 $n/2$ 与 $n-2(6<n)$ 之间至少有一个素数		伯特兰	1845		俄国	切比雪夫	1854	初等方法	肯定		
狄利克雷原理	极小化狄利克雷积分 $\iint\left\{\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right\} dxdy$ 的函数 u 满足位势方程 $\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$	德国	狄利克雷	1851			魏尔斯特拉斯 希尔伯特	1870 1899	举出反例 限制区域边界值，允许函数性质	否定 复活		此猜想是在黎曼的博士论文中公布的。(1851 年)
杰波夫猜想	相邻平方数之间至少存在二个素数		杰波夫	1855	不完全归纳法	 中国	因凡涅斯与品兹 楼世拓姚琦	1983 1984	解析法 筛法	未定 证明 θ 可取 23/42 证明 $\theta>6/11$		

续表 5

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
黎曼猜想	函数 $\xi(s)=\frac{1}{1^5}+\frac{1}{2^5}+\frac{1}{3^5}+\cdots\cdots$（其中 $s=\sigma+t_i$ 为复数）的零点全部落在复平面的一条直线 $\sigma=\frac{1}{2}$ 上	德国	黎曼	1859		 中国	赫克斯雷 莱文生 楼世拓	1973 1974 1980	逐次趋近 逐次逼近	未定 $N(a,T)\leqslant T^{2\cdot4(1-\sigma)}$ $N_0(T)>0.3474N(T)$ 改进了莱文生的结果	如果得到解决，则可以解决一大批数学难题	
连续统假设	在“可数”的势与“连续统”的势之间没有其它的势	德国	康托尔	1882			谢宾斯	1934		未定 列举 12 个等价问题		
高斯猜想	奇数分布函数 $\pi(x)$ 与对数积分函数 $L_i(x)$ 之差为定号，即 $L_i(x)-\pi(x)>0$	德国	高斯	19 世纪			列特雯得	1914		证明了当 n 充分大时，$L_i(x)$ 与 $\pi(x)$ 之差的符号无限改变，从而否定了高斯猜想		
泰特猜想	任何 3-连通的三次平面图都是哈密尔顿的		泰特	19 世纪末			托特	1946 年	举反例	否定		
尺规作图猜想	所有边数等于费尔马数 $F(n)=2^{2^n}+1$ 中素数的正多边形，均可用尺规作图作出来	德国	高斯		不完全归纳法	德国 高斯				肯定 亲自用尺规作出正 17 边形		

续表 6

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
波文猜想	方程 $1^n + 2^n + \cdots + m^n = (m+1)^n$ 只有正整数解 $n = 1, m = 2$		波文							未定		
巴切特猜想	对 $n = 1,2,3,\cdots$，方程 $n^2 = x^2 + y^2 + z^2 + W^2$ 至少有一组 x, y, z, W 的非负整数解		巴切特		不完全归纳法				逐次趋近	未定 当 $n = 1, 2, 3, \cdots, 325$ 时，验证是正确的		
希尔伯特猜想	在等腰三角形中，如果底角与顶角之比是代数数，但不是有理数，则底边与侧边之比总是超越数	德国	希尔伯特	1900	逐次猜想（类比）							
彭加勒猜想	在 n 维空间中的一个点集，若是 $n-1$ 连通的紧致流形，则必定是 n 维球		彭加勒	1904		美国 美国 葡萄牙 英国	斯梅尔 弗里德曼 莱戈 罗克	1960 1981 1986		肯定 证明当 $n \geq 5$ 时成立 证明当 $n = 4$ 时成立 证明当 $n = 3$ 时成立		当 $n = 1, 2$ 时早已被证明成立

续表 7

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
瑞曼伊见猜想	若数 $\tau(n)$ 满足：$\tau(1)x+\tau(2)x^2+\tau(3)x^3+\cdots=x(1-x)^{24}(1-x^2)^{24}(1-x^3)^{24}\cdots$，且 $\tau(p_1^{r1}p_2^{r2}\cdots p_n^{rn})=\tau(p_1^{r1})\tau(p_2^{r2})\cdots\tau(p_n^{rn})$，则有 $\tau(p^2)=\tau(p)^2-p^{11}$，$\tau(p^3)=\tau(p)^3-2p^{11}(p)$，…		瑞曼伊见	1916			摩德尔	1917		肯定	本世纪30年代，赫查推广了这一结果，发展为“模形式理论”这一重要函数论的分支学科	
比巴霸赫猜想	若函数 $f(x)=x+\sum_{n=2}^{\infty}a_nx^n$（$x$ 为复变数）在其定义域单位圆内（$\|x\|<1$）单值连续，且当 $x=0$ 时，有 $f(0)=0$，$f'(0)=1$，则 $\|a_n\|\leqslant n$	德国	比巴霸赫	1916	不完全归纳法	美国	路易斯·德布朗格斯	1984	逐次趋近（算子理论，初等方法）	肯定		
布拉斯奇	任何一个布拉斯奇流形，除一常因数外，必与一个标准布拉斯奇流形等距同胚。		布拉斯奇	1921			格林	1963		肯定	对微分几何、拓扑学、数学分析、近世代数、均有重要价值	

续表 8

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
克拉莫猜想	当 $x = p_n$，$Y = p_n^{1/2}\log p_n$ 时，在区间〔x，$x+y$〕中必定有素数存在		克拉莫	1921	逐级猜想（假定黎曼猜想成立条件下					未定	只要黎曼猜想被证明成立，此猜想即为定理	
莫德尔猜想	任一不可约有理系数的二元多项式，当它的“亏数”大于或等于2时，最多只有有限个解		莫德尔	1922		德国	弗尔廷斯	1984		肯定		
范德瓦尔登猜想	设 A 是 $n \times n$ 矩阵，矩阵元为 $a_{ij}(i=1,2,\cdots,n,j=1,2,\cdots,n)$，则 A 的正项行列式（Permanent）per（A）定义为 $\mathrm{Per}(A) = \sum_{\sigma = s_n} a_{1,\sigma}(1)a_{2,\sigma}(2)\cdots a_{n,\sigma}(n)$，其中 s_n 表示 n 个符号的对称群	德国	范德瓦尔登	1926		苏联	费里克曼 埃戈伊切夫	1979 1980		肯定		

续表 9

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
勒默猜想	不存在合数 n 使得 $\varphi(n) \mid n-1$，即数论函数方程 $k\varphi(n)=n-1(k\geqslant 2)$ 无正整数解		勒默	1932		 中国 中国	勒默 柯召 孙琦 柯召 孙琦	1932 1962 1963	逐次趋近	证明 n 至少有 7 个素数乘积。 证明 n 至少是 12 个素数乘积。 证明 n 至少是 13 个素数乘积。		$\varphi(n)$ 为欧拉函数
超越数猜想	α^{β}，其中底 α 是代数数，指数 β 是代数无理数，例如数 $\alpha^{\sqrt{2}}$ 或 $e^{\pi}=i^{-2i}$ 总是超越数，至少是个无理数	瑞士	欧拉			苏联 德国	盖尔方德 施耐德尔	1934	抽屉原则，丢番图逼近思想等	肯定		人们有时也称其为欧拉猜想
博舒克猜想	是否可能把 E^{d}（d 维欧氏空间）中的每个子集 A 分成 $d+1$ 个子集，其中每个子集的直径都小于 $\mathrm{diam}A$.	波兰	博舒克	1933	不完全归纳法					未定		

续表 10

名　称	内　　容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结　　果		
关于方程的极限环的猜想（Ⅰ）	方程 $x + \mu\sin x + x = 0$ 有无限多个环	美国	Echw-eiler	1946			Hoch-stadt Step-han D'H-eedene	1967 1969	逐次逼进	未　定 证明了当 $\lvert\mu\rvert$ 充分小时，方程在 $\lvert x\rvert \leqslant (n+1)\pi$ 上至少存在 n 个极限环 对于 $\lvert\mu\rvert < 2$ 证明了上述同样的结论，而对 $\lvert\mu\rvert \geqslant 2$ 证明在原点的充分大邻域之外存在无限多个极限环。		
关于方程的极限环的猜想（Ⅱ）	方程 $x + \mu\sin x + x = 0$ 在 $\lvert x\rvert \leqslant (n+1)\pi$ 上恰好存在 n 个极限环	中国	张芷芬	1958		中国	张芷芬	1980		肯　定 证明了在 $\lvert x\rvert \leqslant (\Lambda+1)\pi$ 上恰好存在 n 个极限环		
魏尔猜想			魏尔	1949	类比		德利涅	1974		肯　定		
欧德斯猜想	对于一切 $n>1$ 的正整数，方程 $\frac{4}{n} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}$ 均有正整数解 x, y, z		欧德斯	1950		 中国	史芬斯 柯召 孙琦 张先觉 杨曼托	 1963		未　定 提出猜想：当 $n>2$ 时，此方程的解 x, y, z 满足 $x \neq y, y \neq z, x \neq z$ 但证明当 $2<n<5000$ 时成立。证明欧德斯猜想与史芬斯猜想等价，且证明当 $n<400000$ 时，欧德斯猜想成立。 证明当 $n<10^7$ 时，欧德斯猜想成立。		

续表 11

名　称	内　　容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结　果		
正质量猜　想	当具有洛伦茨度量的四维流形的极大类空超曲面，设其度量张量 $g_{ij} = \left(1+\frac{a}{2r}\right)^4\delta_{ij} + 0\left(\frac{1}{r}\right)$ 时，则表示全质量的常数 a 非负，只当具有欧氏度量时，a 为 0			1959		美国（华裔）	丘成桐 斯科恩	1978 与 1979	非线性编微分方程和变分法	肯　定		
柯召—孙琦猜想	方程 $x^n+(x+1)^n+\cdots+(n+h)^n=(x+h+1)^n$ 除解(1)当 $n=1,h=1$ 时，$x=1$；(2)当 $n=2,h=1$ 时，$x=3$；(3)当 $n=3,h=2$ 时，$x=3$ 以外，无其它正整数解	中国	柯召 孙琦	1962	不完全归纳法	中国	柯召 孙琦 邹兆南	1978		未　定，已证明当 $1 \leqslant n \leqslant 400$ 时，此猜想成立。		
商高数猜　想	对于正整数 a,b,c,x,y,z，如果有 $a^2+b^2=c^2$ 和 $a^x+b^y=c^z$，则 $x=y=z=2$					中国	柯召 孙琦	1963 1964	已证明当 $a=2n+1$，$b=2n(n+1)$，$c=2n(n+1)+1$ 时，猜想对 $n<6144$ 成立	未　定		

续表 12

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
牛曼猜想	任意 n 个连续整数 $m+1, m+2, \cdots, m+n$ 总可以重新排列成 $m+i_1, \cdots, m+i_n$，使 $(m+i_j, j)=1$，$(j=1, \cdots, n)$		牛曼				藏金 贝弥斯	1963		未定 证明了 $n=m$ 时成立。		
						中国	柯召 孙琦	1963		证明了当 $n \leqslant \|70\|6$ 时成立。		
						中国	尹文霖 闵嗣鹤	1963		找出并证明与该猜想等价的命题，并证明当 $n=10$ 和 100 时成立。		
							卡福特	1971		证明了当 $n \leqslant 1002$ 时成立。		
四平方猜想	每个正整数都是四个或四个以下平方数之和									肯定		
场站设置猜想	在一平面上几点连线总长度最小时，其连线间的结点角皆不小于 120°				物理模拟法					未定 已证明当 $n=3$ 时成立，当 $n>3$ 时，条件是必要的，但至今未找到其充分条件		
康托假设												
卡拉毕猜想												

续表 13

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
正规族的 Hayman 猜想	设 $a(a \neq 0)$, b 是两个有穷复数, $n(n \geqslant 5)$ 是正整数。又设 $\mathcal{F}$ 是区域 D 亚纯函数簇,并且对 $\mathcal{F}$ 中每个函数 $f(z)$ 在 D 内有 $f'(z) - af(z)^n \neq b$, 则 $\mathcal{F}$ 在 D 内正规。	英国	Hayman	1964		中国	李先进	1985		肯定		
Smale 猜想	有理 A 微分同胚的 ξ 函数是有理的	美国	Smale	1967		英国	Guckenheimeimer, J. Manning, A. Shub, M.	1970 1971 1978	逐次逼进	肯定		
Lax 猜想	对 KDV 方程 $u_i + 6uu_n + u_{xxx} = 0$ 有如下猜想:存在 N 个正常数 $C_{i,j} = 1,2,\cdots,N$ 和 $2N$ 个常数, $\theta_j^{\pm}$, $j = 1,2,\cdots,N$ 对方程的任一解 $\mu(x,t)$ 有 $\lim\limits_{t\to\pm\infty} u(x+ct,t) = \begin{cases} s(x,\theta_j^{\pm}), \\ 0, \end{cases}$ 当 $C = C_j$ 时 当 $C \neq C_j$ 时 其中, s 为孤立波		Lax, P.D.	1968		中国	徐邦清	1986		肯定		

续表 14

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
临界图猜想	每个临界图均是奇阶的		Jakobsen	1973	不完全归纳法	新加坡	叶宏博	1980		肯定		
Miller 猜想	若 $f(z)\in J(a,0)$, $\mid z\mid = r<1$, 则 $a>0$ 时, $-\frac{\partial}{\partial r}K_0(a,-r)\leqslant\mid f'(z)\mid\leqslant\frac{\partial}{\partial r}K_0(a,r)$, 其中, 设 $a\geqslant 0$, 记单位圆盘 $D=\{z;\mid z\mid<1\}$ 内满足 $f(z)f'(z)/z\neq 0$, $Re\left\{(1-a)\frac{tf'(z)}{f(z)}+a\left(1+\frac{tf''(z)}{f'(z)}\right)\right\}>0$ 的正则函数 $f(z)=z+a_2z^2+\cdots$ 的全体记为 $J(a,0)$. $K_0(a,z)=\left[\frac{1}{2}\int_0^2\xi^{1/a-1}(1-\xi)^{-2/a}d\xi\right]^a$	美国	Miller	1973		中国	马万仓	1986	特例法	肯定		
费利普和斯透特猜想	当 $t\to+\infty$ 时, $\forall\delta>0$ 有 $s(t)-X(t/2)\langle\langle t^{a+\delta}a\cdot s$, 其中 $X(t)$ 为正则布朗运动, $s(t)$ 与 $\sum_{k\leqslant\tau}\cos 2\pi n_k w$ 同分布, 费利普和斯透特证明了 $a=5/12$, 猜想 a 可降为 1/3.	美国	费利普斯透特	1975	直观推断	中国	孙志刚	1982		肯定		1984 年邵满进一步把 a 降为 1/4.

续表 15

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
Szymiczek 猜测	存在两个域分别适合下面两组性质： A.i. 平方类数 $[F:F^{*2}]=8$(其中 F^* 表 F 之乘群)ii. 二次型 x^2+y^2(F 中)一切数，且存在另一个与 x^2+y^2 不等价之二元二次型表一切数 iii. 存在一个不表一切数之二元型。 B.i. 平方类数 $=8$ ii. 该域是形式实的，且恰有一种方法定序，且每一正数都是二平方数之和 iii. 该域中唯一表一切数(精确到等价)的二元二次型 x^2-y^2		Szymiczek	1975		中国	李德琅	1980		肯定 构造出无穷多个满足 Szymiczek 猜测的两组条件的子域。		
时滞方程猜测	时滞微分方程 $x(t)=-x^{\frac{1}{3}}(t)+x^{\frac{1}{3}}(t-r)$ 的各个解当 $t\to\infty$ 时趋于常数		Bernfeld Haddock	1976	直观推断	中国	丁同仁	1981		肯定		
码恒交换等价于前缀码猜想	码恒交换等价于前缀码		Perrin Schützenberger	1977		中国	章亮 郭未琦 王利民	1985	等价命题法	未定 转化为与它等价的两个命题之一		

续表 16

名称	内容	提出状况				研究进展					意义	备注
		国别	姓名	时间	方法	国别	姓名	时间	途径	结果		
等部有向图猜想	对于形如 $DK_n(m)$ 的变分有向图族，可分条件是存在同构因子分解的充分条件	美国 澳大利亚 澳大利亚	Hararr Robins-on Worma-ld	1978	不完全归纳法	中国	王建方	1983	逐次逼进	肯定		
何成奇猜想	设 $w=f(z)$ 是 $\|z\|<1$ 内的 K- 拟共形映照，就范条件为 $f(0)$、$f(re^{i\frac{2k\pi}{n}})=r^{\frac{1}{k}}e^{i\frac{2k\pi}{n}}$，$0\leqslant k\leqslant n-1$，（*）其中 n 是正整数，r 为(0,1)内任一实数。猜想：对于适合就范条件（*）的 K- 拟共形映照必存在等角分布的 n 条射线，映照最大星形域复盖这些射线之长的算术平均值将无限地接近于 $4^{\frac{\pi k}{nk}}$.	中国	何成奇	1981	直观推断	中国	何成奇	1984		肯定		
Moran 猜想	若 $x_1,x_2,\cdots x_n,\cdots$ 为相互独立随机变数序列，且有相同的分布 $p(x_i=1)=p(x_i=-1)=\frac{1}{2}$ 那么对任何实数 $a_1\geqslant a_2\geqslant\cdots\geqslant a_n\geqslant 0$，$n\geqslant 1$，可能成立不等式 $E\left\|\sum_{i=1}^{n}a_ix_i\right\|\geqslant\frac{1}{\sqrt{2}}\left(\sum_{1}^{n}a_i^2\right)^{\frac{1}{2}}$	澳大利亚	Moran			中国	刘坤会	1984	逐次逼进	肯定		

[General Information]
书名=数学猜想集
作者=徐本顺　解恩泽
页数=253
SS号=10252913
DX号=
出版日期=1998年05月第2版
出版社=湖南科学技术出版社

封面
书名
版权
前言
目录
一、源远流长--从勾股定理到费尔马大定理
(一) 从色股三角形谈起
(二) 勾股数
(三) 问题的拓广与特例
1．由“变”到“常”，由“常”到“变”
2．由相同到相异
3．由多到少，由少到多
4．由特殊到一般，由一般到特殊
5．由形到数，由数到形
6．从数的性质提出问题
7．由类比提出问题
(四) 一条著名的旁注
二、长路漫漫--费尔马大定理的探讨
(一) n=4的费尔马大定理
(二) 关于n=3的欧拉证明
1．欧拉关于n=3的证明
2．根式环
3．关于两平方数之和
4．一个引理的证明
5．关于两个平方数之和的注记
6．欧拉证明的基本思路
(三) 关于n=3的一个初等证明
(四) 从勒让德到库姆尔
1．关于n=5和n=7，分圆整数
2．代数数论基本知识
3．关于正则素数
4．其它一些结果

(五) 费尔马大定理研究的一些新成果

1．考虑结论反面的必要条件

2．充分条件

(六) 简评

三、触类旁通--费尔马大定理与莫德尔猜想

(一) 莫德尔猜想

(二) 解不定方程的一般性问题

(三) 几个重要结果

31．曲线的沙发列维奇(shafarevich)猜想

2．阿贝尔簇的沙发列维奇猜想

3．有界高度原理

4．同源下高的行为

5．泰特猜想

(四) 莫德尔猜想的证明

(五) 从莫德尔猜想到费尔马大定理

(六) 模曲线和费尔马大定理

(七) 费尔马大定理获证之后

四、一步之遥？--哥德巴赫猜想

(一) 猜想的提出

(二) 悲观的预言与惊人的成果

(三) 圆法

(四) 筛法

五、补天何须五色石--地图着色与四色猜想

(一) 四色猜想的提出

1．什么叫四色猜想

2．先生问学生和学生问先生

(二) 早期的证明和五色定理

1．凯利的呼吁

2．另辟蹊径

3．约当曲线和欧拉定理

4．五色定理

5．肯普的证明

(三) 四色猜想的证明
1．不可避免组和可约构形
2．公开宣称的一种信念
3．等价的形式
4．可约性障碍和放电
5．新的困难
6．人机合作证明了四色猜想
7．解决地图四色问题的重大意义
(四) 平面图
(五) 线(边)着色
(六) 顶点着色
(七) 全色猜想
六、法无定法--提出数学猜想的若干方法
(一) 不完全归纳法
(二) 类比法
(三) 变换条件法
(四) 物理模拟法
(五) 联系观察法
(六) 逐级猜想法
七、闪光的并非都是金子--判定数学猜想真伪性的几个途径
(一) 举例否定
(二) 逐次趋近
(三) 命题转化
(四) 反证法
八、千淘万漉始到金--数学猜想的艰难性
(一) 有一个逐步完善的过程
(二) 时间长与途径曲折
1．时间长
2．途径曲折
(三) 有时得不到多数人的承认
九、数学猜想的类型、特征与意义
(一) 数学猜想的类型

1．存在型猜想
2．规律型猜想
3．方法型猜想
(二) 数学猜想的特征
1．真伪的特定性
2．思想的创新性
3．目标的具体性
(三) 研讨数学猜想的重要意义
1．丰富数学理论
2．促进数学方法论的研究
3．推动潜科学学的探讨